

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 13:04:36 : Search time 1693.29 seconds
(without alignments)
10396.748 Million cell updates/sec

Title: US-10-019-783-3

Perfect score: 10966
Sequence: 1 ctcgacccatctgcatgtc.....cagacagcctggcgccgc 10966

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	96.2	0.9	1926	US-09-249-585A-4	Sequence 4, Appl
C 2	96.2	0.9	1931	US-09-130-114-2	Sequence 2, Appl
C 3	95.4	0.9	390	US-09-197-649-7	Sequence 7, Appl
C 4	95.4	0.9	4403765	US-09-103-840A-2	Sequence 2, Appl
C 5	94	0.9	4411529	US-09-103-840A-1	Sequence 1, Appl
C 6	93.8	0.9	4411529	US-09-103-840A-1	Sequence 1, Appl
C 7	91.4	0.8	4403765	US-09-103-840A-2	Sequence 2, Appl
C 8	87.4	0.8	612	US-09-902-540-1357	Sequence 1357, Ap
C 9	78.6	0.7	985	US-09-056-556-182	Sequence 182, App
C 10	78.6	0.7	985	US-09-072-596-177	Sequence 177, App
C 11	78.6	0.7	985	US-09-072-596-177	Sequence 182, App
C 12	77	0.7	2434	US-09-489-847-67	Sequence 67, Appl
C 13	75.6	0.7	1039	US-09-902-540-1280	Sequence 1280, Ap
C 14	75.4	0.7	114793	US-10-148-806-3	Sequence 3, Appl
C 15	72.4	0.7	1141	US-09-806-708B-22	Sequence 22, Appl
C 16	72	0.7	2277	US-08-676-967-5	Sequence 5, Appl
C 17	72	0.7	2277	US-08-676-967-5	Sequence 5, Appl
C 18	72	0.7	2277	US-09-058-487-5	Sequence 5, Appl
C 19	71.8	0.7	1736	US-09-182-816-22	Sequence 22, Appl
C 20	71.8	0.7	1736	US-09-182-816-24	Sequence 24, Appl
C 21	71.8	0.7	1736	US-09-471-528-22	Sequence 22, Appl
C 22	71.8	0.7	1736	US-09-471-528-24	Sequence 24, Appl
C 23	71.8	0.7	1736	US-09-634-530-22	Sequence 22, Appl
C 24	71.8	0.7	1736	US-09-634-530-24	Sequence 24, Appl
C 25	71.2	0.6	2481	US-09-894-998A-35	Sequence 35, Appl
C 26	71.2	0.6	2481	US-10-237-551-35	Sequence 35, Appl
C 27	71.2	0.6	3066	US-10-237-551-152	Sequence 152, App

28	71.2	0.6	154746	4	US-09-827-688-8	Sequence 8, Appl
C 29	71.2	0.6	154746	4	US-09-827-688-8	Sequence 8, Appl
C 30	71.2	0.6	396	4	US-09-640-173-53	Sequence 53, Appl
C 31	71.2	0.6	396	4	US-09-713-550-53	Sequence 53, Appl
C 32	71.2	0.6	396	4	US-09-825-294-53	Sequence 53, Appl
C 33	71.2	0.6	396	4	US-09-970-966-53	Sequence 53, Appl
C 34	70	0.6	1141	4	US-09-806-708B-22	Sequence 22, Appl
C 35	70	0.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 36	69.6	0.6	7218	1	US-08-232-463-14	Sequence 14, Appl
C 37	69.4	0.6	601	4	US-09-949-016-30530	Sequence 30530, A
C 38	69.4	0.6	601	4	US-09-949-016-30531	Sequence 30531, A
C 39	69.4	0.6	601	4	US-09-949-016-37149	Sequence 37149, A
C 40	69.4	0.6	601	4	US-09-949-016-37150	Sequence 37150, A
C 41	69.4	0.6	601	4	US-09-949-016-37153	Sequence 37153, A
C 42	69.4	0.6	601	4	US-09-949-016-37163	Sequence 37163, A
C 43	69.4	0.6	601	4	US-09-949-016-145867	Sequence 145867, A
C 44	69.4	0.6	601	4	US-09-949-016-145868	Sequence 145868, A
C 45	69.4	0.6	601	4	US-09-949-016-146135	Sequence 146135, A

ALIGNMENTS

RESULT 1
US-09-249-585A-4/c
Sequence 4, Application US/09249585A
Patent No. 6417002
GENERAL INFORMATION:
APPLICANT: Horlick, Robert
TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISODES
FILE REFERENCE: 0667/0D905
CURRENT APPLICATION NUMBER: US/09/249,585A
CURRENT FILING DATE: 1999-02-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 1926
TYPE: DNA
ORGANISM: Epstein Barr Virus
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)..(1926)
OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match 0.9%; Score 96.2; DB 3; Length 1926;
Best Local Similarity 49.2%; Pred. No. 1.1e-11;
Matches 282; Conservative 0; Mismatches 288; Indels 3; Gaps 1;

QY	666	AGAGCGACGAGTGTCCCGGACGCGCTTCCGTCGCGACGCGGACGCGAAGACA	725
DB	855	AGAGCGGAGAGAGACGAGACGCGGAGACGCGGAGAGACGAGACGCGGAGACG	796
QY	726	ACGCGCATGCGTGTCCGCGCTGTAAACGCAAGACCAACGCGCATGCGTGTATCCG	785
DB	795	GGAAGAGACGAGACGCGGAGAGACGCGGAGACGAGACGCGGAGAGACGAG	736
QY	786	ACGCGAAGCGCAAGACGAGACGCGCATGCGTGTCCGACGCGAAGCGCAAGCAACG	845
DB	735	ACGCGGAGAGACGAGACGCGGAGAGACGAGACGCGGAGAGACGAGACGCGG	676
QY	846	GCAATGCGGAGCGCATGCGGACGCGGACGCGGACGCGGACGCGGACGCGGACG	902
DB	675	AGAGCGGAGAGAGACGAGACGCGGAGACGAGACGCGGAGAGACGCGGAGAG	616
QY	903	ACGCGCACCGAGACGAGACGCGCATGAGCGCGCGACGCGAAGCGGAGAGACG	962
DB	615	AGAGCGAGAGACGCGGAGAGACGAGACGCGGAGAGACGCGGAGAGAGAGAG	556
QY	963	AGCATGCGGAGACTCCGCGCGGACGCGGAGAGACGCGGAGAGACGCGGAGAG	1022
DB	555	AGAGCGGAGAGACGCGGAGAGAGACGAGACGCGGAGAGAGACGAGAGACG	496

QY 1023 AGAGAGAGAGAGCGCTGTGAGTGTGATTTCGGGGTCCCAAGAGACGGCTGTGCGCGCA 1082
DB 495 GGGAGGAGCGGGGAG 436
QY 1083 CGGGGGGAGACATGAGAGATCCGGGCGATTACGTTCAAGATCAGGCGCAAGCGTTCAGAGAGA 1142
DB 435 AG 376
QY 1143 AGGGGCGCGCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
DB 375 AGGACGGGGAG 316
QY 1203 GCACGCGCGCTGAGAGCGCGAG 1235
DB 315 ACAGAGAGCGGGGAG 283

RESULT 2

US-09-130-114-2/c
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassem B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Bplisomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

Query Match 0.9%; Score 96.2; DB 2; Length 1931;
Best Local Similarity 49.2%; Pred. No. 1.1e-11;
Matches 283; Conservative 0; Mismatches 288; Indels 3; Gaps 1;

QY 666 AGAGCGAGAGAGTCCCGCGAG 725
DB 855 AGAGCGGGGAG 796
QY 726 AGCGGCGAG 785
DB 795 GGGAG 736
QY 786 AGCGGAG 845
DB 735 AGCGGGGAG 676
QY 846 GCGATGCGGAG 902
DB 675 AGAGCGGGGAG 616
QY 903 AGCGGCGAG 962
DB 615 AGAGCGAG 556
QY 963 AGCATGCGGAG 1022
DB 555 AGAGCGGGGAG 496
QY 1023 AGGAG 1082
DB 495 GGGAG 436
QY 1083 CGGGGGGAGACATGAGAGATCCGGGCGATTACGTTCAAGATCAGGCGCAAGCGTTCAGAGAGA 1142
DB 435 AG 376

QY 1143 AGGGGCGCGCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
DB 375 AGGACGGGGAG 316
QY 1203 GCACGCGCGCTGAGAGCGCGAG 1235
DB 315 ACAGAGAGCGGGGAG 283

RESULT 3

US-09-197-649-7
; Sequence 7, Application US/09197649
; Patent No. 6194550
; GENERAL INFORMATION:
; APPLICANT: Gold, Larry
; APPLICANT: Tuerk, Craig
; APPLICANT: Pridnow, David
; APPLICANT: Smith, Jonathan D.
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
; FILE REFERENCE: NEX02/CI-CON
; CURRENT APPLICATION NUMBER: US/09/197,649
; CURRENT FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: 07/829,461
; EARLIER FILING DATE: 1992-01-31
; EARLIER APPLICATION NUMBER: 07/739,055
; EARLIER FILING DATE: 1991-08-01
; EARLIER APPLICATION NUMBER: 07/561,968
; EARLIER FILING DATE: 1990-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
; OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-197-649-7

Query Match 0.9%; Score 95.4; DB 3; Length 390;
Best Local Similarity 53.3%; Pred. No. 5.3e-12;
Matches 201; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 670 CGAGGAGTCCCGCGAG 729
DB 14 CGAGGAG 73
QY 730 CCATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789
DB 74 CGAGGAG 133
QY 790 GAAAGGAG 849
DB 134 CGAGGAG 193
QY 850 TGCCGAGGCGACTGCGAG 909
DB 194 CGAGGAG 253
QY 910 CGCGGAG 969
DB 254 CGAGGAG 313
QY 970 CGAGGAG 1029
DB 314 CGAGGAG 373
QY 1030 GGAAGGCGTGTGAGTGA 1046
DB 374 GGTTCGATGTTCAGGA 390


```
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FRASER, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
```

```
Query Match 0.9%; Score 93.8; DB 3; Length 441529;
Best Local Similarity 49.4%; Pred. No. 9.9e-09;
Matches 272; Conservative 0; Mismatches 277; Indels 2; Gaps 1;
```

```
QY 659 GTACGCCAGAGCGAGTCCGCGAAGCGCCCTTGCCGCGAGCCGCGAAGCGC 718
DB 675266 GTCCGCGCGCGCGCGCGCGCGAGCAACCGCATCGCGGTGTACGGGTACTCGGCC 675207
QY 719 AAGAGCAACCGCCATGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 778
DB 675206 AGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 675147
QY 779 GATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 838
DB 675146 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 675087
QY 839 AGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 898
DB 675086 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 675027
QY 899 ACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 958
DB 675026 GCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 674967
QY 959 AACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1018
DB 674966 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 674909
QY 1019 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078
DB 674908 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 674849
QY 1079 GCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1138
DB 674848 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674789
QY 1139 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1198
DB 674788 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 674729
QY 1199 TTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1209
DB 674728 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 674718
```

```
RESULT 7
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FRASER, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
Query Match 0.8%; Score 91.4; DB 3; Length 4403765;
Best Local Similarity 46.6%; Pred. No. 3.6e-08;
Matches 293; Conservative 0; Mismatches 336; Indels 0; Gaps 0;
```

```
QY 663 GCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 722
DB 3934692 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3934751
QY 723 GCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 782
DB 3934752 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3934811
QY 783 CCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 842
DB 3934812 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3934871
QY 843 ACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 902
DB 3934872 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3934931
QY 903 ACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 962
DB 3934932 CCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3934991
QY 963 AGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1022
DB 3934992 CAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935051
QY 1023 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1082
DB 3935052 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935111
QY 1083 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1142
DB 3935112 ACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935171
QY 1143 AGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1202
DB 3935172 TGAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935231
QY 1203 GCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1262
DB 3935232 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935291
QY 1263 GCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1291
DB 3935292 GCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935320
```

```
RESULT 8
US-09-902-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 683447
; GENERAL INFORMATION:
```


APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(115849)B
CURRENT FILING DATE: 2001-07-10
PRIORITY APPLICATION NUMBER: 60/217,883
PRIORITY FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1357
LENGTH: 612
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(612)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 0.8%; Score 87.4; DB 4; Length 612;
Best Local Similarity 50.1%; Pred. No. 5.4e-10;
Matches 214; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

QY 4598 AACAAATTTTAAACAGCAATTTTGAATTTGACAAATTTTAAACGGGTAT 4657
DB 143 AAAT 202
QY 4658 TCCTGACATTTTCAAAATTTGATCAAAATTTTAAACGACTTTCTCAATTTGA 4717
DB 203 AATTA 262
QY 4718 GCATATTTTAAATTTAAAGTTCACAAATTTGCACTTTTAAATTTAGCGAGAA 4777
DB 263 AATAAAATTTAAATAATTTAAATAATTTAAATAATTTAAATAATTTAAATA 322
QY 4778 CATTTGAAATTTTAAATTTTGAATTTGCAATTTTCTATTTCTGACAAAT 4837
DB 323 AA 382
QY 4838 TGAATAATGCAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 4897
DB 383 AATAAAATTTAAATAATTTAAATAATTTAAATAATTTAAATAATTTAAATAATTA 442
QY 4898 ACATATTTTGAATAACAAATTTTAAAGTAAATTTAAATTTAAATTTAGAA 4957
DB 443 AAAAAAAAAATTTAAATAATTTAAATAATTTAAATAATTTAAATAATTTAAATA 502
QY 4958 ACATTAATAATGCAAAATTTAAAGAAATCCGAGAAAGCACTGGAATGCAAT 5017
DB 503 AATTA 562
QY 5018 GGAATAA 5024
DB 563 AAAAAAA 569

RESULT 9
US-09-056-556-182
Sequence 182, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yael A.W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 182:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-056-556-182

Query Match 0.7%; Score 78.6; DB 3; Length 985;
Best Local Similarity 48.5%; Pred. No. 8.8e-08;
Matches 311; Conservative 0; Mismatches 319; Indels 11; Gaps 3;

QY 649 CATTGCCACCGTACCGACGACGAGTCCGCGCAACGCGCTTGGCGCGCAGC 708
DB 3 CAGCGCTACCGGTGCGCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 62
QY 709 CGCGACGCGCAAGCAAGCGCCATGCGGTGCGCGCGCGCGCGCGCGCGCGCGCG 768
DB 63 CAACAGCGGTGCG 122
QY 769 CCATGCGGTGATCCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 828
DB 123 CGGACCGGAGGTCCG 182
QY 829 GAACGCGCAAGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 888
DB 183 TGGCG 239
QY 889 GAACGCGCAAGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948
DB 240 TACCG 299
QY 949 CGCGGAGCGCAAGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1002
DB 300 CGGCG 359
QY 1003 -GCATGCG 1061
DB 360 TGAACG 419
QY 1062 AGAAGCGCGGTGCG 1121
DB 420 CGGCG 479
QY 1122 TCAGCGCGAGGTGCG 1181
DB 480 AGGTCTGCGCGCAACG 538
QY 1182 CGTCTGTTTCCCG 1241
DB 539 ACGGCGCGACGCA 598
QY 1242 TGCGACCG 1282
DB 599 ATGCGCGCAACG 639

RESULT 10
US-09-072-596-177
Sequence 177, Application US/09072596
Patent No. 6458366
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedic, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 177:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-596-177

Query Match 0.7%; Score 78.6; DB 3; Length 985;
Best Local Similarity 48.5%; Pred. No. 8.8e-08;
Matches 311; Conservative 0; Mismatches 319; Indels 11; Gaps 3;

QY 649 CATGGCCACCGCTAGCCAGACGAGTGGCCGCGGACCGCTTGGCCGCGGCGACG 708
DB 3 CAGGCGTACCGGTGGCGCGGCGGTTCGCGGCGCGCGCGGAGAGCGCGGCGG 62
QY 709 CGCGAAGCGCAAGAGCAAGCGCCATGCGGTGCTGCGCGCGCTGAACGCGCAAGAGCAAGCG 768
DB 63 CAACAGCGGTGTGGCGCGGCAACAGCGTCCGCGCGCGCGCGCGCGGTGACGCGGCAAGCG 122
QY 769 CCATGGCGTGGATGCCGACGCGGAAAGCGGAGAGCAAGCGCCATGCGGTGCTGCGGACG 828
DB 123 CGGCAACGCGAGGTCCGCGCGGTCCGCGCGGAGCAACCGCCGCTGGGTTCGCGCGG 182
QY 829 GAACGGCAAGAGCAAGCGCCATGCGGAGCGCACTGCGGCAAGCGCGCGGCGCACTGC 888
DB 183 TGGCGCGCGCGGCAAGG---TGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 239
QY 889 GAACGGCAAGAGCAAGCGCCACCGCGAGAGCAAGCGCCATGCGGTGCGCGCGGAGCA 948
DB 240 TACCGCGCGGCAACGCGCGCGGTTCGCGCGGCAACCGGTAGTGCAGCATCGCGCGGCGCGG 299

QY 949 CGCGAGAGCAAGCATGCGCGGAGACTCCGCGCGGAGAGCGGAGCAAGCG----- 1002
DB 300 CGCGCGCGCGGAGTGAACGCGCGGCGATGGGGCGAGCGGTCTGGGCTTGGGCTT 359
QY 1003 -GCATGCGCGCGCGCGCGGCAAGAGAGAGAGAGCGCGGTGAGTGAATTTCGCGGCTGCA 1061
DB 360 TGAACGCGCGGCAAGCGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCATCA 419
QY 1062 AGACGCGCGTGTGCGCGGCGGAGCGGCGGAGCAATGACATCCGCGCGATACGATCAAGA 1121
DB 420 CGGCGCGCGCGCGCGCGCGCGGCAACGCGCGGAGAGCGCGGAGAGCGCGCGCGCG 479
QY 1122 TCAGCGGAGCGGTGACAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGGAGC 1181
DB 480 AGGTCTCGCGGAGCAAGCGCGGCGGTGCGCGGTGACGCGGTGAGGCGCGGTGCGGCGC -CGCGGCA 538
QY 1182 CGTCCGTGTTCCCGGCTTCCGCAAGCGCGGTGAGGCGCGGAGAGCGCGGTGCGCGCGCGC 1241
DB 539 ACAGCGGCAAGCGCGGCGGTGCGCGGTGACAGCAAGCGCGGAGAGCGCGCGCGCGGCA 598
QY 1242 TGGCAGCGGCGCAAGTCACTGCTAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1282
DB 599 ATGCGGCAAGCGGCG 639

RESULT 11
US-09-072-967-182
Sequence 182, Application US/09072967
Patent No. 6592877
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedic, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
NUMBER OF SEQUENCES: 355
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,967
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 182:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-967-182

Query Match 0.7%; Score 78.6; DB 4; Length 985;
Best Local Similarity 48.5%; Pred. No. 8.8e-09;
Matches 311; Conservative 0; Mismatches 319; Indels 11; Gaps 3;

QY 649 CATGGCAACCGTACCGGAGAGCGAGTCCGCGGAAACGCGCTTGGCGTGGCCGAGC 708
Db 3 CAGGCTACCGGTGGCCCGCGGTTCCCGCGCGCGCGGAGAGCGAGCGGCGG 62
QY 709 CGCGAAGCGCAAGCAAGCGGCAATGGGTGGTCTGGCGCGTGAACGCAAGAGCAAG 768
Db 63 CACAGCGGTGTGGCGCGGCAACAGCGCTCCGCGCGCGCGCGGTGACGCGGCAAGG 122
QY 769 CCATGGCGTGGATCGCGGAGCGGCAAGCGGCAAGCGGCGGCGTGGCGGAGCGC 828
Db 123 CGGACCGGAGGTCCGCGGCGGTCCGCGCGGAGCAACCGGCTGGTTCGCGCG 182
QY 829 GAACGGCAAGCAAGCGGCAATGGCGGAGCGGCAAGCGGCGGCGGCAAGCGC 888
Db 183 TGGCGCGCGGCGGCAAGG---TGGCGCGCGCGCGCGCGCGCGCGCGCGGCGGCGG 239
QY 889 GAACGGCAAGCAAGCGGCAATGGCGGAGCGGCAAGCGGCGGCGGCAAGCGC 948
Db 240 TACCGCGGCGGCAAGCGGCGGCGTGTTCGCGCGGCAAGCGGCGGCGGCGG 299
QY 949 CGGCGAGAGCAAGCGGCAATGGCGGAGCGGCAAGCGGCGGCGGCAAGCGG----- 1002
Db 300 CGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 359
QY 1003 -GCATGCGCGCGCGGCGGCGGCGGAGAGAGAGCGGCGGCGGCGGCGGCGGCGG 1061
Db 360 TGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 419
QY 1062 AGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1121
Db 420 CGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 479
QY 1122 TCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1181
Db 480 AGGTCTCGGCGCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 538
QY 1182 CGTCTGTGTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1241
Db 539 ACGCGGCGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 598
QY 1242 TGGCGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1282
Db 599 ATGGCGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 639

RESULT 12
US-09-489-847-67
Sequence 67, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489, 847
EARLIER FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 67
LENGTH: 2434
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (12)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (27)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (73)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (75)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (103)
OTHER INFORMATION: n equals a,t,c,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (130)
OTHER INFORMATION: n equals a,t,c,g, or c
US-09-489-847-67

Query Match 0.7%; Score 77; DB 4; Length 2434;
Best Local Similarity 54.4%; Pred. No. 4e-07;
Matches 178; Conservative 0; Mismatches 145; Indels 4; Gaps 1;

QY 4661 TGAACATTTTCAAAATTTGATCAAAATTTTAAAGACCTCTTCTCAATTTGAGCA 4720
Db 2110 TGATCATTTTACAGAAATTTATGAAAGATTTTGAATTTTGTGATGCT 2169
QY 4721 ATATTTAAATTTATTAATAAGTTCAACATTTTGAATTTTAAATTTAGCGAAGAT 4780
Db 2170 ATGTTAGCGGAAACCAATTTTATGATTTTAAACCTGATGAGAAACATTTGACA 2229
QY 4781 TTGAATTTCTAAATTTTGAATTTGAAATTTTCTATTTCTGAACAAATTTGA 4840
Db 2230 TGTATATGCTCACTTTCTCAATTTTCTGCTAATTTTCTAAGATACATTTAAATGTT 2289
QY 4841 AATATCAAGCTAATTTGATTAATTTTGAATAATGCAATTTTGAATTTTGAACA 4900
Db 2290 TTATAT---TTTTTTTAAATTAATGACCGGTAAGAAATTAATTAATTAATTAATTA 2345
QY 4901 TATTTGAAAAAACAATACTTTAAAGTTAAATTAATTAATTAATTAATTAATTAATTA 4960
Db 2346 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2405
QY 4961 TAAAAATTAAGCAAAAAATAAAGAAA 4987
Db 2406 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2432

RESULT 13
US-09-902-540-1280
Sequence 1280, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

Query Match	0.78;	Score 75.6;	DB 4;	Length 1039;
Best Local Similarity	48.3%;	Pred. No. 4.6e-07;		
Matches 210; Conservative	0;	Mismatches 225;	Indels 0;	Gaps 0;

Qy	4590	AAATTGGCAACAAATTAAAAAAGCGAACAATTTTGAATTGGACCAATTTTAA	4649
Db	66	AACTTAAAAACAAACCAAAATTTAAAAAATTTAAAAAACAATTTATATAA	125
Qy	4650	ACGGGATTCCTGCAACATTTTCAAATGTGCATCAAAATTTAAACGACTCTTCTC	4709
Db	126	AAAAAAATAAAAACAAATATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA	185
Qy	4710	AAATTGAGCAATTTTAAAAATTAATTAATAAAAAAGTCAACAATTTTGAATTTTAAAAAT	4769
Db	186	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCTTTATTAATAAAAAAAAAAAAAAAAAATATATA	245
Qy	4770	AGCGAGAACATTTTGAATTTCTAAATATTTTGCATTTTGAACAATTTTTCATTTTCGA	4829
Db	246	AAAAAAAAACATACATATAAAAAAAAAAAAAAAAAAATATAATATAAAAAATCAAAAAAAAAA	305
Qy	4830	ACAAAAATGTGAAATACGACGTATTTTGGATTAATTTTGGAAAAATCCGATTTTTCGA	4889
Db	306	AAAAATAAAAAATAAAAAAAAAATTTAAATCAATCATATAAAAAAAAAATTTAAATTAATACACA	365
Qy	4890	ATTTCTGACATATTTTGAATAAAACAAAAAATCTTAAAGATAATTAATTAATTAATTA	4949
Db	366	AAAAAAACAAAAATTTAAAAATTAATAAATTTAATTTACACAAATTTAAAGAAAAATTAATAA	425
Qy	4950	AAATAGAAACATAAAATTAAGCAAAAAAATTAAGAAATTCGAGAAAAATGCCAATCGGAA	5009
Db	426	AAAAAAAAAAAAATATATGTAAAAAATAAAACAAACAAAAACACCAAAAAAAAAAAAAATATAA	485
Qy	5010	TACGACATGAAAAA 5024	
Db	486	TAAAAAATAAAAAA 500	

```

RESULT 14
US-10-148-806-3/c
Sequence 3, Application US/10148806
Patent No. 6762042
GENERAL INFORMATION:
APPLICANT: Bai, Chang
APPLICANT: Metzger, Michael
APPLICANT: Liu, Xiaomei
TITLE OF INVENTION: DNA MOLECULES ENCODING
FILE REFERENCE: 20585P
CURRENT APPLICATION NUMBER: US/10/148,806
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: US00/33065
PRIOR FILING DATE: 2000-12-09
PRIOR APPLICATION NUMBER: 60/169,970
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq
SEQ ID NO 3

```

Query Match	0.74	Score 75.4	DB 4	Length 114793
Best Local Similarity	48.24	Pred. No. 1.5e-05		
Matches 274	Conservative 0	Mismatches 291	Indels 4	Gaps 2

QY	673	CGAGGTGGCCGCCA	CGGCCTTTCGCTGGGCCGACGCCGGAACCGGCAAGGCAACGGGCCA	732
Db	31497	CGGCTGGCGGAAAGTGTCCGGGGGCGGGGCGGGGCGGAGCTCGAGCGGGCAGGGGCGG	31438	
QY	733	TGGCGTGGCTGCCGCGTGAACGGCAAGACCAACGGCCATGGCGTGGATGCGGACCGGA	792	
Db	31437	GGGCGGGGGCCGGGGCGAAGCTCGGAGCGGGCAAGGGCCGGGGCGGGGCGGGAGCT	31378	
QY	793	CGGCAAGACCAACGGCCATGGCGTGGCTGCGGACCGGAACCGGCAAGCAACGGCCATGC	852	
Db	31377	CGGAGCGGGGCGGGGCGGGGCGGGGCGGGGCGGAGCTCGAGCGGGGCGGGGCGGGG	31318	
QY	853	CGAGGGCACTGGCGGACGGGGCCAGGGGCACTGGGAAACGGCAACCAACGGGCCACGG	912	
Db	31317	CGGGGGGAGGGCTGA---GCAAGCTGAGGGGGGGGGGCACTGGGCGGGGGCGGGGCGAG	31267	
QY	913	CGAGAGCAACGGCCATGCTGAGGGCGGCCGACCGCAACGGCCAGGACGACATGCGCA	972	
Db	31260	GGGAGGCTCGGAGCGGGGCAAGGGGCGGGGCGGGGCGGGGCGGAGCTCGAGGCGGGCAAG	31201	
QY	973	GGACTCCGGCGGCGAACGGCGAGAGCAACGGGCACTGCGGCGGGCGGCGGCAAGAGGAGGA	1032	
Db	31200	GGCGGGGGCGG--GGGCAAGGCTGTGACGACGAGCGGGGGCGGGGCACTGGGCGGGGGCGG	31142	
QY	1033	GGCGGTGAGGTGAATTTGCGGGGTTGCCAAGGACGGCGTGGCTGGCGGCGGCGGCGGA	1092	
Db	31141	GGCGAAGGCGGAGCTCTGAGACGGGGCGAGGGGCGGGGCGGGGCCAGGGCGGAGCTCGAGGC	31082	
QY	1093	CATGACATTCGGGCGCATACGATACAAATCAACGCGGACGTTGCAAGAAAGGGCGGG	1152	
Db	31081	GGGCGAAGGGGCGGGGGGGGGGGGCAAGGCTTAGCAGCAGAGGGGCGGGGCACTGGGCGGG	31022	
QY	1153	GGCGGTGCTGCGCGTGGCCCAAGGGGACCGTCCGTGTTCCGGGCTTCCGCAAGGCGGT	1212	
Db	31021	GGCGGGGCTTAGGGCTGAGCTCGGAGGGGGCGAGAGCGGGGCGGGGACTTCGAGGCGCTG	30962	
QY	1213	CGAGGCGGAGAACGCGGTGCGCGGCGCGC	1241	
Db	30961	AGCTGAGGTGAGGGTGGGGGGCTCTCTCC	30933	

```

RESULT 15
US-09-806-708B-22/c
: Sequence 22, Application US/09806708B
: Patent No. 6784342
: GENERAL INFORMATION:
: APPLICANT: The University of British Columbia
: TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
: FILE REFERENCE: 4810-58741
: CURRENT APPLICATION NUMBER: US/09/806,708B
: CURRENT FILING DATE: 2001-04-03
: PRIOR APPLICATION NUMBER: US 60/147,133
: PRIOR FILING DATE: 1999-08-04
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 22
: LENGTH: 1141
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: NAME/KEY: promoter
: LOCATION: (1)..(1141)
: OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

```


C 21	17.6	73.3	276337	4	US-09-944-076-17504
C 22	17.2	71.7	285	4	US-09-927-767-191
C 23	17.2	71.7	285	4	US-09-920-767-15473
C 24	17.2	71.7	417	4	US-09-513-959C-2780
C 25	17.2	71.7	1938	4	US-10-196-927-1
C 26	17.2	71.7	2765	4	US-09-799-451-50
C 27	17.2	71.7	2847	4	US-10-196-927-5
C 28	17.2	71.7	4200	1	US-07-841-654B-1
C 29	17.2	71.7	4200	1	US-07-946-234A-1
C 30	17.2	71.7	4200	1	US-08-123-161A-1
C 31	17.2	71.7	4200	1	US-08-483-278-1
C 32	17.2	71.7	4200	5	PCP-US93-01560-1
C 33	17.2	71.7	2433	4	US-09-639-207-9
C 34	17.2	71.7	177937	4	US-09-944-016-4123
C 35	17.2	71.7	227979	4	US-09-944-016-11842
C 36	16.8	70.0	2265	4	US-09-949-039A-442
C 37	16.8	70.0	5933	4	US-09-918-172-43
C 38	16.6	69.2	601	4	US-09-944-016-7716
C 39	16.6	69.2	601	4	US-09-944-016-1823
C 40	16.6	69.2	1132	4	US-09-270-767-769
C 41	16.6	69.2	1132	4	US-09-270-767-16051
C 42	16.6	69.2	1247	3	US-09-947-035-1
C 43	16.6	69.2	2005	1	US-07-916-991-5
C 44	16.6	69.2	2165	1	US-08-222-289-1
C 45	16.6	69.2	19736	3	US-09-740-035-3

Sequence 17504, A
Sequence 191, App
Sequence 15473, A
Sequence 2780, Ap
Sequence 1, Appl
Sequence 50, Appl
Sequence 5, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 9, Appl
Sequence 1425, A
Sequence 11842, A
Sequence 442, Appl
Sequence 23, Appl
Sequence 27515, A
Sequence 18235, A
Sequence 769, App
Sequence 16051, A
Sequence 1, Appl
Sequence 5, Appl
Sequence 1, Appl
Sequence 3, Appl

ALIGNMENTS

```

RESULT 1
US-09-799-451-184/C
Sequence 184, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Auerndt, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aifeng J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yundong
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Raddeje T.
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: fl_Fl_genes Version 2.0
SEQ ID NO 184
LENGTH: 2081
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (28)..(1938)
US-09-799-451-184

```

Query Match	78.3%	Score 18.8;	DB 4;	Length 2081;
Best Local Similarity	90.9%;	Pred. No. 16;		
Matches	20;	Conservative	2;	Indels 0;
		Mismatches		Gaps 0

QY 2 CTGGATCTCAGTGCATCCAC 23
|||||
Db 1011 CTGGGTTCTCAGTGCATCCAGC 990

```

RESULT 2.21
US-09-544-398B-11/c
Sequence 11, Application US/09544398B
Patent No. 6770461
GENERAL INFORMATION:
APPLICANT: Carulli, John P.
APPLICANT: Little, Randall D.
APPLICANT: Becker, Robert R.
APPLICANT: Johnson, Mark L.
TITLE OF INVENTION: High bone mass gene of 11q13.3.
FILE REFERENCE: 032796-013
CURRENT APPLICATION NUMBER: US/09/544,398B
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/229,319
PRIOR FILING DATE: 1999-01-13
PRIOR APPLICATION NUMBER: US 60/071,449
PRIOR FILING DATE: 1998-01-13
PRIOR APPLICATION NUMBER: US 60/105,511
PRIOR FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 641
SOFTWARE: Paetssq for Windows Version 4.0
SEQ ID NO: 11
LENGTH: 66933
TYPE: DNA
ORGANISM: Homo sapiens
US-09-544-398B-11
Query Match 75.8%; Score 18.2; DB 4;
Best Match Similarity 87.0%; Pred. No. 56;
Matches 20; Conservative 0; Mismatches 3;
OY 2 CTGGGATCTCGTGCATCCACACA 24
DB CTGGGAGCACAGTCCACACACA 15733

```

; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-544-398B-11

```

RESULT 3 3871B-11/c
Sequence 11, Application US/09543771B
US-09-543-771B-11/c
Patent No. 6780609
GENERAL INFORMATION:
APPLICANT: Carrulli, John P.
APPLICANT: Little, Randall D.
APPLICANT: Becker, Robert R.
APPLICANT: Johnson, Mark L.
TITLE OF INVENTION: High bone mass gene of 11q13.3.
FILE REFERENCE: 034796-014
CURRENT APPLICATION NUMBER: US/09/543, 771B
CURRENT FILING DATE: 2000-04-05
PRIORITY APPLICATION NUMBER: US 09/229,319
PRIORITY FILING DATE: 1999-01-13
PRIORITY APPLICATION NUMBER: US 60/071,449
PRIORITY FILING DATE: 1998-01-13
PRIORITY APPLICATION NUMBER: US 60/105,511
PRIORITY FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 641
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 66933
TYPE: DNA
ORGANISM: Homo sapiens
US-09-543-771B-11

```

Query Match	75.8%;	Score 18.2;	DB 4;	Length 66933;
Best Local Similarity	87.0%;	Pred. No. 56;		
Matches	20;	Conservative	0;	Mismatches
			3;	Indels
				Gaps
0x	2	CTGGGATTCAGTGCATCCACAA	24	

Db 15755 CTGGAGACAGTCGACGACCA 15733

RESULT 4

US-09-544-3988-9/c
Sequence 9, Application US/095443988
Patent No. 6770461
GENERAL INFORMATION:
APPLICANT: Canull, John P.
APPLICANT: Little, Randall D.
APPLICANT: Recker, Robert R.
APPLICANT: Johnson, Mark L.
TITLE OF INVENTION: High bone mass gene of 11q13.3
FILE REFERENCE: 032796-013
CURRENT APPLICATION NUMBER: US/09/544,3988
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/229,319
PRIOR FILING DATE: 1999-01-13
PRIOR APPLICATION NUMBER: US 60/071,449
PRIOR FILING DATE: 1998-01-13
PRIOR APPLICATION NUMBER: US 60/105,511
PRIOR FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 641
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 72049
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (8356), (8385), (38585)
OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-544-3988-9

Query Match 75.8%; Score 18.2; DB 4; Length 72049;
Best Local Similarity 87.0%; Pred. No. 56;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGGATCTCAGTCGATCCACA 24

Db 19346 CTGGAGACAGTCGACGACCA 19324

RESULT 5

US-09-543-771B-9/c
Sequence 9, Application US/09543771B
Patent No. 6780609
GENERAL INFORMATION:
APPLICANT: Canull, John P.
APPLICANT: Little, Randall D.
APPLICANT: Recker, Robert R.
APPLICANT: Johnson, Mark L.
TITLE OF INVENTION: High bone mass gene of 11q13.3
FILE REFERENCE: 032796-014
CURRENT APPLICATION NUMBER: US/09/543,771B
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 09/229,319
PRIOR FILING DATE: 1999-01-13
PRIOR APPLICATION NUMBER: US 60/071,449
PRIOR FILING DATE: 1998-01-13
PRIOR APPLICATION NUMBER: US 60/105,511
PRIOR FILING DATE: 1998-10-23
NUMBER OF SEQ ID NOS: 641
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 72049
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (8356), (8385), (38585)
OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno

US-09-543-771B-9

Query Match 75.8%; Score 18.2; DB 4; Length 72049;
Best Local Similarity 87.0%; Pred. No. 56;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGGATCTCAGTCGATCCACA 24

Db 19346 CTGGAGACAGTCGACGACCA 19324

RESULT 6

US-09-949-016-14199/c
Sequence 14199, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14199
LENGTH: 140844
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(140844)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14199

Query Match 75.8%; Score 18.2; DB 4; Length 140844;
Best Local Similarity 87.0%; Pred. No. 63;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGGATCTCAGTCGATCCACA 24

Db 87157 CTGGAGACAGTCGACGACCA 87135

RESULT 7

US-08-422-699A-8/c
Sequence 8, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Houman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746
US-08-422-699A-8

Query Match 74.2%; Score 17.8; DB 2; Length 2511;
Best Local Similarity 90.5%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGAGTTCAGTCGATCCACA 24
DB 2407 GGGGTCTCAGTCGATCCAAA 2387

RESULT 8
US-08-422-706B-8/c
Sequence 8, Application US/08422706B
Patent No. 597733
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millicia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,706B
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/284,543
FILING DATE: 08-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746
US-08-422-706B-8

Query Match 74.2%; Score 17.8; DB 2; Length 2511;
Best Local Similarity 90.5%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGAGTTCAGTCGATCCACA 24
DB 2407 GGGGTCTCAGTCGATCCAAA 2387

RESULT 9
US-08-422-699A-12/c
Sequence 12, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millicia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-422-699A-12

Query Match 74.2%; Score 17.8; DB 2; Length 2726;
Best Local Similarity 90.5%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTCGATCCAA 24
DB 2660 GGGGCTCAGTCGATCCAAA 2640

RESULT 10
US-08-422-706B-12/c
Sequence 12, Application US/08422706B
Patent No. 5977333
GENERAL INFORMATION:
APPLICANT: Brookman, J. David
APPLICANT: Houseman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,706B
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/284,543
FILING DATE: 08-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-422-706B-12

Query Match 74.2%; Score 17.8; DB 2; Length 2726;
Best Local Similarity 90.5%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTCGATCCAA 24
DB 2660 GGGGCTCAGTCGATCCAAA 2640

RESULT 11
US-08-484-044-11/c
Sequence 11, Application US/08484044
Patent No. 5552282
GENERAL INFORMATION:
APPLICANT: Caskey, C. T.
APPLICANT: Fu, Ying-Hui
APPLICANT: Friedman, David L.
APPLICANT: Pizzuti, Antonio
APPLICANT: Fenwick, Raymond G.
TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pulbright & Jaworski, L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,044
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/019,940
FILING DATE: 19-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5443
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 3182 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-484-044-11

Query Match 74.2%; Score 17.8; DB 1; Length 3182;
Best Local Similarity 90.5%; Pred. No. 53;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTGCATCCACA 24
DB 3104 GGGGTCTCAGTGCATCCAAA 3084

RESULT 12
US-08-422-699A-10/C
Sequence 10, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545

FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-9540
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 3323 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(518..3323, "")
OTHER INFORMATION: /standard_name="cDNA 41"
FEATURE:
NAME/KEY: misc_difference
LOCATION: replace(769..3323, "")
OTHER INFORMATION: /standard_name="cDNA 28"
US-08-422-699A-10

Query Match 74.2%; Score 17.8; DB 2; Length 3323;
Best Local Similarity 90.5%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGGATCTCAGTGCATCCACA 24
DB 3219 GGGGTCTCAGTGCATCCAAA 3199

RESULT 13
US-08-422-706B-10/C
Sequence 10, Application US/08422706B
Patent No. 5977333
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,706B
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/264,543
FILING DATE: 08-AUG-1994
PRIOR APPLICATION DATA:

```

1 APPLICATION NUMBER: US 08/023,612
2 FILING DATE: 26-FEB-1993
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: US 07/839,255
5 FILING DATE: 20-FEB-1992
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: PCT/US93/01545
8 FILING DATE: 19-FEB-1993
9 PRIOR APPLICATION DATA:
10 APPLICATION NUMBER: PCT/GB93/00253
11 FILING DATE: 05-FEB-1993
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: GB9202485.0
14 FILING DATE: 06-FEB-1992
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Grahman, Patricia
17 REGISTRATION NUMBER: 32,227
18 REFERENCE/DOCKET NUMBER: MIT-5830A2
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: 617-861-6240
21 TELEFAX: 617-861-9540
22 INFORMATION FOR SEQ ID NO: 10:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 3323 base pairs
25 TYPE: nucleic acid
26 STRANDEDNESS: double
27 TOPOLOGY: linear
28 MOLECULE TYPE: DNA (genomic)
29 FEATURE:
30 NAME/KEY: misc_difference
31 LOCATION: replace(518..3323, "")
32 OTHER INFORMATION: /standard_name= "CDNA 41"
33 FEATURE:
34 NAME/KEY: misc_difference
35 LOCATION: replace(769..3323, "")
36 OTHER INFORMATION: /standard_name= "CDNA 28"
37
38 US-08-422-706B-10
39
40 Query Match 74.2% Score 17.8; DB 2; Length 3323;
41 Best Local Similarity 90.5%; Pred. No. 54;
42 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
43
44 QY 4 GGGATCTCAGTCATCCACA 24
45 ||| ||||| ||||| |||||
46 Db 3219 GGGGTCTCAGTCATCCAAA 3199
47
48 RESULT 14
49 US-08-484-044-10/C
50 Sequence 10, Application US/08484044
51 Patent No. 5552282
52 GENERAL INFORMATION:
53 APPLICANT: Caskey, C. T.
54 APPLICANT: Pu, Ying-Hui
55 APPLICANT: Friedman, David L.
56 APPLICANT: Pizutti, Antonio
57 APPLICANT: Fenwick, Raymond G.
58 TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy
59 NUMBER OF SEQUENCES: 13
60 CORRESPONDENCE ADDRESS:
61 ADDRESSEE: Fulbright & Jaworski, L.L.P.
62 STREET: 1301 McKinney, Suite 5100
63 CITY: Houston
64 STATE: Texas
65 COUNTRY: U.S.A.
66 ZIP: 77010-3095
67
68 COMPUTER READABLE FORM:
69 MEDIUM TYPE: floppy disk
70 COMPUTER: IBM PC compatible
71 OPERATING SYSTEM: PC-DOS/MS-DOS
72 SOFTWARE: PatentIn Release #1.0, Version #1.25
73 CURRENT APPLICATION DATA:
74 APPLICATION NUMBER: US/08/484.044

```

```

1  FILING DATE:
2  CLASSIFICATION: 435
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER: US 08/019,940
5  FILING DATE: 19-FEB-1993
6  ATTORNEY/AGENT INFORMATION:
7  NAME: Paul, Thomas D.
8  REGISTRATION NUMBER: 32,714
9  REFERENCE/DOCKET NUMBER: D-5443
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: 713/651-5325
12 TELEFAX: 713/651-5246
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014

```

Query Match	74.2%	Score 17.8	DB 1	Length 11613
Best Local Similarity	90.5%	Pred. No. 66		
Matches 19	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	4	GGGATCTGATGTCATCCACCA	24	
DB	11107	GGGGTCTGATGTCATCCAAAA	11087	

Oy 4 GGATCTCAGTCATCCACA 24
Db 1107 GGGTCTCAGTCATCCAAA 11087

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 73.3%; Score 17.6; DB 1; Length 1722;
 Best Local Similarity 83.3%; Pred. No. 61;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCTGGGATCTCAATGCATCCACA 24
 Db 1122 TTGGGATGCCAATGCATCCACA 1145

Search completed: October 15, 2005, 03:25:12
 Job time : 5.70591 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 13:04:36 / Search time 30422.4 seconds
(without alignments)
13720.565 Million cell updates/sec

Title: US-10-019-783-3
Perfect score: 10966
Sequence: 1 ctgcgtccatcgcacatgctc.....cagacagcctcggcgccgc 10966

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :

EST:
1: gb_ests1:
2: gb_ests2:
3: gb_hic:
4: gb_ests3:
5: gb_ests4:
6: gb_ests5:
7: gb_ests6:
8: gb_ests7:
9: gb_ests8:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	545.4	5.0	681	4	BI956132 HVSMEM002
2	461.8	4.2	657	6	CA015631 HT14015T
3	439.6	4.0	770	2	BF268125 HV CBA001
4	384.2	3.5	692	6	CD870152 AZO2.113J
5	383	3.5	730	6	CD864826 AZO2.001M
6	376	3.4	592	6	CD879356 AZO4.104P
7	363.8	3.3	980	6	CC692577 OG0J39TV
8	363.4	3.3	606	5	BQ66157 HT01121T
9	359.6	3.2	610	7	CN009140 WHE3835_C
10	345	3.1	354	6	CA002831 HS08K11F
11	340.4	3.1	373	3	AJ433473 AJ433473
12	340	3.1	635	6	CD878050 AZO4.101M
13	337	3.1	360	1	AJ485409 AJ485409
14	325.6	3.0	940	9	CG033764 PUFYB66TB
15	316	2.9	934	9	CC696106 OG0J015TV
16	312.4	2.8	980	9	CG343192 OG0AL69TV
17	311.8	2.8	481	2	BF200495 WHE2258_G
18	309.2	2.8	498	2	BE490536 WHE0362_D
19	306	2.8	694	6	CD866252 AZO2.102P
20	301.6	2.8	704	4	BJ269947 BJ269947
21	301.2	2.7	689	6	CD878493 AZO4.102O
22	300.4	2.7	555	5	BQ661382 HM03B06u
23	300	2.7	300	1	AJ485410 AJ485410
24	297.6	2.7	886	9	CG033766 PUFYB66TB

25	279.6	2.5	473	5	BQ469047	BQ469047 HM03B06r
26	273.2	2.5	917	4	BM100620	BM100620 EBma01_SO
27	267.6	2.4	407	2	BF260044	BF260044 HVSMEM002
28	266	2.4	402	2	BE587818	BE587818 WHE063_A
29	259.6	2.4	588	6	CA608149	CA608149 wr1.PK008
30	255.8	2.3	732	4	BJ274931	BJ274931 BJ274931
31	254.8	2.3	733	6	CD866253	CD866253 AZO2.102P
32	253.4	2.3	1335	9	CL966192	CL966192 OG1PC013
33	252.4	2.3	486	1	AL816989	AL816989 AL816989
34	251.8	2.3	592	6	CD864994	CD864994 AZO2.073A
35	251.8	2.3	657	4	BJ275175	BJ275175 BJ275175
36	251.4	2.3	806	6	CD439867	CD439867 ELO1N0530
37	244.8	2.2	320	6	CA652974	CA652974 wre1n.pk1
38	244.8	2.2	565	2	BF065437	BF065437 HV CBA001
39	244.4	2.2	371	6	CA000437	CA000437 HS08K11N
40	243.4	2.2	820	8	BZ641440	BZ641440 OGCK047C
41	243.4	2.2	829	8	BZ534170	BZ534170 OGAKH17TM
42	243.4	2.2	923	9	CG368887	CG368887 OG0GP95TV
43	242.2	2.2	740	6	CB646006	CB646006 OSJNB08C
44	241.4	2.2	510	5	BU985095	BU985095 HF06B21T
45	240.6	2.2	745	6	CB678552	CB678552 OSJNB01L

ALIGNMENTS

RESULT 1
BI956132
LOCUS
DEFINITION
HVSMEM0025M09F Hordeum vulgare green seedling EST library
HVCNDA0014 (Blumeria infected) Hordeum vulgare subsp. vulgare cDNA
clone HVSMEM0025M09F, mRNA sequence.

ACCESSION
BI956132.1 GI:16303042
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.

REFERENCES

Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Chin, A., Begum, D.,
Fritsch, D., Atkins, M., Yu, Y., Henry, D., Palmer, M., Rambo, T.,
Simmons, J., Gates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomes: Blumeria infected Morex (compatible) seedling
cDNA library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 402
Seq primer: AATTAACCTCCTAAGG
High quality sequence stop: 597.
Location/Qualifiers

JOURNAL COMMENT

FEATURES
- source
1..681
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/culivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMEM0025M09F"
/tissue_type="green seedling leaf"
/lab_host="TJ0121"
/clone_lib="Hordeum vulgare green seedling EST library
HVCNDA0014 (Blumeria infected)"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Morex (mla) plants were greenhouse grown in the R
Wise lab at Iowa State University, Ames, IA; 7 day old

QY 1717 CGACATGCACTGCTGATCCATGCGCAGAAACACCGCATGTCATCATAA 1776
 DB 290 CGACATGCACTGCTGATCCATGCGCAGAAACACCGCATGTCATCATAA 349
 QY 1777 CCCCAACAACCGCTGCGGAGGCTTACTCTCTACGACCATCTGTCAGATTTCATCC 1836
 DB 350 CCCCAACAACCGCTGCGGAGGCTTACTCTCTACGACCATCTGTCAGATTTCATCC 396
 QY 1837 TTGGCTTGCTGAATATGATTCAGTTGACGTGACCTGTCGAATCTTTTGGCAATGCC 1896
 DB 397 ----- 396
 QY 1897 ATACTGACTGATGTTGCTCAATTAAGTCCGCGAGGTGCGAAAGGCTCGAATATGCT 1956
 DB 397 -----AGTCCGAGAGTGGGAAAGGCTCGAATATGCT 433
 QY 1957 GATTGCTGACGAGTATACGCGAAGTGTCTTGGGACGCGCCGTTTCATCCATGCG 2016
 DB 434 GATTGCTGACGAGTATACGCGAAGTGTCTTGGGACGCGCCGTTTCATCCATGCG 493
 QY 2017 AGTGTGGGCAATCAACCCGCTGCTGCTCCATAGGCTCTCTGTCGAATCATGATAGT 2076
 DB 494 AGTGTGGGCAATCAACCCGCTGCTGCTCCATAGGCTCTCTGTCGAATCATGATAGT 553
 QY 2077 GCGTGGATGCGCGCTTGATGAGTGGTACGCGGTATACGACCCGAGAAAGATCTTACAGAAAC 2136
 DB 554 GCGTGGATGCGCGCTTGATGAGTGGTACGCGGTATACGACCCGAGAAAGATCTTACAGAAAC 613
 QY 2137 TAAAGTACTTAATC 2151
 DB 614 TAAAGTACTTACATC 628

RESULT 3
 LOCUS BF268125 770 bp mRNA linear EST 23-OCT-2001
 DEFINITION HV_CEA0019P10F Hordeum vulgare seedling green leaf EST library
 HVCNA0004 (Blumeria challenged) Hordeum vulgare subsp. vulgare
 cDNA clone HV_CEA0019P10F, mRNA sequence.

ACCESSION BF268125 GI:13263792
 VERSION BF268125.2 GI:13263792
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 770)
 Wing, R., Close, T.J., Kleinof, A., Wise, R., Wei, F., Begum, D.,
 Fritsch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J.,
 Choi, D.W., Fenton, R.D., Oates, R. and Main, D.
 Development of a genetically and physically anchored EST resource
 for barley genomics: Blumeria infected incompatible (Mla13)
 seedling leaf cDNA library
 Unpublished (2001)
 On Nov 17, 2000 this sequence version replaced gi:11199120.

JOURNAL
 COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 586
 Seq primer: AATTAACCTCACTAAAGG
 High quality sequence stop: 611.
 Location/Qualifiers

FEATURES
 source
 1..770
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="C16155 (Mla13)"
 /sub_species="vulgare"
 /db_xref="taxon:112509"

ORIGIN

Query Match 4.0%; Score 439.6; DB 2; Length 770;
 Best Local Similarity 83.6%; Pred. No. 2e-78;
 Matches 565; Conservative 0; Mismatches 14; Indels 97; Gaps 2;

/clone="HV_CEA0019P10F"
 /issue_type="seedling green leaf"
 /lab_host="TUC121"
 /clone_lib="Hordeum vulgare seedling green leaf EST
 library HVCNA0004 (Blumeria challenged)"
 /note="Vector: lambda2AP; Site 1: BcoRI; Site 2: XhoI;
 C.I. 16155 (Mla13) plants were greenhouse grown in the R
 Wise lab at Iowa State University, Ames, IA; 7 day old
 green seedlings were challenged with isolate A27
 (AvrMla13) of Blumeria graminis f. sp. hordei, and leaves
 were harvested 20 and 24 hr post-inoculation and snap
 frozen; uninoculated leaves were harvested 20 hr
 post-inoculation (Wei, Wise). In the TJ Close lab at the
 University of California, Riverside, total RNA was
 prepared from each sample pool, equal quantities of all
 three RNA pools were combined, poly(A) RNA was purified
 from the mixture, one cDNA library was made, and 1 million
 pfu were in vivo excised to give Bluescript SK(-) cDNA
 phagemids (Choi, Close). Phagemids were plated and picked
 at the Clemson University Genomics Institute (CUGI)
 (Begum, Palmer, Fritsch, Atkins and Wing). Plasmid DNA
 preparations, DNA sequencing and sequence analysis were
 performed at CUGI (Wing, Yu, Fritsch, Henry, Simmons,
 Oates, Rambo, Main). The sequence has been trimmed to
 remove vector sequence and contains a minimum of 100 bases
 of phred value 20 or above. For more details on library
 preparation and sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see Close TJ, Wing R, Kleinof A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.ipw.usda.gov/gspages/bgn/31/cover.html)"

QY 1477 ACAAGCGCCGCTGGAGACACCTTCCGAGGCGCTCCGTCATGTCGCGGAGCA 1536
 DB 101 ACAAGCGCCGCTGGAGACACCTTCCGAGGCGCTCCGTCATGTCGCGGAGCA 160
 QY 1537 GCTTCTTCAACCGCGCGGAGACCCAGGCGATGAGGTCATTAATCCGGTCTGAGCCA 1596
 DB 161 GCTTCTTCAACCGCGCGGAGACCCAGGCGATGAGGTCATTAATCCGGTCTGAGCCA 220
 QY 1597 GACCGCGGCGCAATTCCTGCTCCGAGGCGAGCTTACCCAACTAGAGCGCGCG 1656
 DB 221 GACCGCGGCGCAATTCCTGCTCCGAGGCGAGCTTACCCAACTAGAGCGCGCG 280
 QY 1657 GCGTTCACACGCTGAGGTCGCGCATTTGACTCTACCGCAAGAGGGTGGAGAT 1716
 DB 281 GCGTTCACACGCTGAGGTCGCGCATTTGACTCTACCGCAAGAGGGTGGAGAT 340
 QY 1717 CGACATGCACTGCTGATCCATGCGCAGAAACACCGCATGTCATCATAA 1776
 DB 341 CGACATGCACTGCTGATCCATGCGCAGAAACACCGCATGTCATCATAA 400
 QY 1777 CCCCAACAACCGCTGCGGAGGCTTACTCTCTACGACCATCTGTCAGATTTCATCC 1836
 DB 401 CCCCAACAACCGCTGCGGAGGCTTACTCTCTACGACCATCTGTCAGATTTCATCC 447
 QY 1837 TTGGCTTGCTGAATATGATTCAGTTGACGTGACCTGTCGAATCTTTTGGCAATGCC 1896
 DB 448 ----- 447
 QY 1897 ATACTGACTGATGTTGCTCAATTAAGTCCGCGAGGTGCGAAAGGCTCGAATATGCT 1956
 DB 448 -----AGTCCGAGAGTGGGAAAGGCTCGAATATGCT 484
 QY 1957 GATTGCTGACGAGTATACGCGAAGTGTCTTGGGACGCGCCGTTTCATCCATGCG 2016

Query Match	Similarity	Score	DB	Length	980;
Best Local	Similarity	69.1%	Pred.	No.	5.9e-63;
Matches	577;	Conservative	0;	Mismatches	237;
				Indels	21;
				Gaps	5;
7015	ATTAACAAACAGGCGCCGTGACGAGCACTTGTGACAGGGCGTCCCTACAGCTATCG	7014			
166	AAAAAAAAAACAAGCTCTCTGCGGAGCACTGTGACGACCTTCCATACAAAGCTGTGCG	225			
7075	GCCGACGACGCTTCTCTCAACGCGCGGGGAACTCAAGGAGTCAAGTCAATATCCCGGTG	7134			
226	AGCGACGACATCTTGTGTACCGCGCGGAGACGCAAGCACTGAGAGTGTGTCTCACTC	285			
7135	CTGGCCCACTGCGCGCGCAACATACTGCTTCCCGGCGAGGCTATCAAAATTACGAG	7194			
286	CTGGCGAG---CGGGGACCAACATATCTGCTCCCGAGCGGGGTATTCGAACACGAG	342			
7195	GCGGAGCGGCAATTCACAAAGCTGAGGTCCGGCACTTCGACCTCATCCCGACAAAGGG	7254			
343	GCGGCGCGGAGGCTGACCAACCTGGAAATTGCGCGGCTTCAATCTGATCCCGAGAGAGG	402			
7255	TGGGAGATTCGACATCTGCTGCTGGAATTCATTCGCGCGCAAGAAACACACCGCATGGTC	7314			
403	TGGGAGATTCACATTCGACGCTGTGGAGTCGATGCGCGCAAGAAACACACCGCATGGTC	462			

Db	Accession	Source	LOCUS DEFINITION	Version	Keywords	Organism	Reference Authors	Title Journal Comment	Features
QY	7315	ATCATAAACCCAAACAAATCCGTTGGGGGAGCGCTTTAACTCTTACGACATCTGGCCAAAGTT	606 bp	EST 30-MAY-2002		Hordeum vulgare subsp. vulgare	Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner, A.	Barley ESTs from germinating seeds Unpublished (2002) Contact: Stein Nils Molecular Markers Group, Department Genbank Institute of Plant Genetics and Crop Plant Research (IPK) Corrensstr. 3, 06466, Gatersleben, Germany Tel: 039482-5522 Fax: 039482-5595 Email: stein@ipk-gatersleben.de Insert length: 606 Sfd Error: 0.00 Plate: 1 row: L column: 21 Seq primer: T3.	Location/Qualifiers 1. .606 /organism="Hordeum vulgare subsp. vulgare" /mol_type="mRNA" /culivar="barke" /sub_species="vulgare" /db_xref="taxon:112509"
QY	463	ATCATAAACCCCAACACACCTTGGGGAGTGTCTACACCCGAGACATTTGGCCAAAGCA	522			Hordeum vulgare subsp. vulgare			
QY	7375	TTGCATCCATCATCTCTGCTGCTGTTATGACGCGGTCTGTGTAACATAGTATATGGA	743			Hordeum vulgare subsp. vulgare			
Db	523	CGGCTAGCTGCTT-----ATCTTGCTGGCCATTGTGGATTGGATGGATCGG	571			Hordeum vulgare subsp. vulgare			
QY	7435	TTGCGTTTGTCTAATCGTGTCTGATGATGCTGTGTTGGTTATC-AGTTCGGGAGGTGGCA	749			Hordeum vulgare subsp. vulgare			
Db	572	ATCGGTTGTGTGTCTTAAGACACTACTGTCTGTGATCTAGTGTGGAGGTGGCA	631			Hordeum vulgare subsp. vulgare			
QY	7494	AGGAAGTCGGAATATTGGTATGATCGCTGACAGAGTTTACGGCAAACTGGTCTGCGGACG	7553			Hordeum vulgare subsp. vulgare			
Db	632	AGGAAGCTGGAAATACGTGTATCGCTGATAGGTGTACGGAATCTGTGTTGGGAGC	691			Hordeum vulgare subsp. vulgare			
QY	7554	GCCCCGTTTATCCGATGGGCGGTCTTTGGGACATTTGCCCGCGTCTTGTCATTTGGATCT	7613			Hordeum vulgare subsp. vulgare			
Db	692	ACCCCTTTGTCGCCCATGGGTGTCTTTGGCCACATTTGCCCTGTGTTGACCATATGATCA	751			Hordeum vulgare subsp. vulgare			
QY	7614	CTGTCCAAAGTCGTGATATATGTCGCGTGGATATGGGACTTGGATGGGTGGCGGTATACACCC	7673			Hordeum vulgare subsp. vulgare			
Db	752	CTGTCCAAAGATGATATATGTCGCGGTGGGCGACTGTGTGGGTGGCCGTATGATATCC	811			Hordeum vulgare subsp. vulgare			
QY	7674	ACAAAGATTTTATAGAGAAAACCTAAGTAGCTTTAGCTCCCTATCATATCTTCTCATATGCTA	7733			Hordeum vulgare subsp. vulgare			
Db	812	AACAAGTTCTGCAGAAACCAAGTA---ATATATCATTTAGCTCATATCTTATGTCTT	868			Hordeum vulgare subsp. vulgare			
QY	7734	CTGTGGGGATTAGTATTTTCTTAAATTTGTTACTGCTTGTGTTATTCAATCTCTACGT	7793			Hordeum vulgare subsp. vulgare			
Db	869	GTTCACAGTAGTGTGTGCTAAGCT---GTTTCACATCATGATTTTTTGGAGATATCATTTGAT	925			Hordeum vulgare subsp. vulgare			
QY	7794	CTATTAGAAATTACTCTTAATGCTCAACGCGACCGACCACTTGTTCAGGTTAG	7848			Hordeum vulgare subsp. vulgare			
Db	926	CGATCAACAACTTCCTTACCTTTCAACATCCGACATCCGACATTTTGTTCAGGTTAG	980			Hordeum vulgare subsp. vulgare			

Db	225	GAGCAACGGGCA	CGGCGGAGGCGCTGCCGCAACGGCCGAGCAACGGCCATGAGGTGGC	284
Oy	1015	GCGGCGAGAGC-----	AGAGAGAGGCGGTGAGTGGAAATTTCCGCGGTGCCAAGCGG	1066
Db	285	GCGCGCACTGGCGGCGCGCCGAGAGAGCGCGTGGATGGAAATTTCCGCGGCGCCAAAGGCGG		344
Oy	1069	CGTGTGCGCGGCA	CGGGGGCGAAATCATGACATCCGGGCGAATGAGGTCAAGATCAAGGCG	1122
Db	345	CGTCTTGGCGCGACGCGGGCGCAATAGACATCCGGGCGATACGCTACAGATCAAGCGC		404
Oy	1129	GAGCGTCAGAGAGAA	GGGGCGCGCGCCGTGTGCGCTGTGGCCCA	CGGGGACCCGTCGGT 1188
Db	405	GAGCGTCAGAGAGAA	CGGGCGCGCGCCCGTGTGCGCCCGCCCA	CGGGGACCCGTCGGT 464
Oy	1189	GTTCCCGGCTTCCGCA	CGGCGCTGAGGCGCAAGACGCGTGGCGCGCGCGCTGCGCAC	1244
Db	465	TTTCCCGGCTTCCGCA	CGGCGCGCTGAGGCGCAAGACGCGTGGCGCGCGCGCTGCGCAC	524
Oy	1249	CGGCGAGTTCAACTGCTA	CCCCCGCGGCGGTGCGGCTTCCCGGCGCGACGAG	1299
Db	525	GCGGCGAGTTCAACTGCTA	CCCCCGCGGCGGTGCGGCTTCCCGGCGCGACGAG	575

RESULT 10					
CA002831					
LOCUS	CA002831	354 bp	mRNA	linear	EST 23-OCT-2000
DEFINITION	HS08K11r HS <i>Hordeum vulgare</i> subsp. <i>vulgare</i> cDNA clone HS08K11				
	5-PRIME, mRNA sequence.				

ACCESSION	CA002831	GI:24279813
VERSION	CA002831.1	
KEYWORDS	EST.	
SOURCE	Hordeum vulgare subsp. vulgare	
ORGANISM	Hordeum vulgare subsp. vulgare	

REFERENCE
AUTHORS

1 (bases 1 to 354)
Zhang, H., Potockina, E., Michalek, W., Stein, N. and Graner, A.

TITLE	Barley ESTs from germinating seeds
JOURNAL	Unpublished (2002)
COMMENT	Contact: Stein Nils

Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: steinlaupj-gatersleben.de
Insert Length: 354 Std Error: 0.00
Plate: 8 Row: K Column: 11
Seq primer: M13rev.

FEATURES	Location/Qualifiers
source	1. .354

```

/organism="Hordelum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="NCBI:258040"
/db_xref="taxon:112509"
/clone="HS08K11"
/tissue_type="embryo + scutellum"
/dev_stage="0-16 hours after imbibition"
/lab_host="X110-Gold"
/clone_1ib="HS"
/note="vector: pBluescript SK+, Site_1: EcoRI (5'-end of cDNA), Site_2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamI, SalI, PstI). NOTE: Also due to the cloning system used, Blue/white selection for recombinants is not 100% reliable."

```

ORIGIN			
Query Match	3.1%	Score 345;	DB 6; Length 354;
Best Local Similarity	100.0%	Pred. No. 3.3e-59;	
Matches 345; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;

QY	817	GGCTGCGCAGCGCGAACGGCAAGAGCAACGGCTCATGCGAGGCCACTGCTGAAACGGCTACCGG	876
Db	10	GGCTGCGCAGCGCGAACGGCAAGAGCAACGGCTCATGCGAGGCCACTGCTGAAACGGCTACCGG	69
QY	877	CGAGGCCACTGCTGAAACGGCAAGAGCAACGGCTCATGCGAGGCCACTGCTGAAACGGCTACCGG	936
Db	70	CGAGGCCACTGCTGAAACGGCAAGAGCAACGGCTCATGCGAGGCCACTGCTGAAACGGCTACCGG	129
QY	937	CGCGCAGCGCAACGGCGAGAGCAACGAGCATGCGAGGACTCTCGCGCGCAACGGCGAGAG	996
Db	130	CGCGCAGCGCAACGGCGAGAGCAACGAGCATGCGAGGACTCTCGCGCGCAACGGCGAGAG	189
QY	997	CAACGGGCTATGCGCGCGCGCGCGAGAGAGAGAGCGCGGTGAGTGAATTTTCGCGGG	1056
Db	190	CAACGGGCTATGCGCGCGCGCGCGAGAGAGAGAGCGCGGTGAGTGAATTTTCGCGGG	249
QY	1057	TGCGCAAGAGCGGCGTGTCTGCGCGCGAGCGGGGCGAATGAGCATTCGGGCGGATTCGGGTA	1116
Db	250	TGCGCAAGAGCGGCGTGTCTGCGCGCGAGCGGGGCGAATGAGCATTCGGGCGGATTCGGGTA	309
QY	1117	CAAGTCAGCGCGAGCGTGTGAGAGAGAGGGGCCGCGGCCGCTGCT	1161
Db	310	CAAGTCAGCGCGAGCGTGTGAGAGAGAGGGGCCGCGGCCGCTGCT	354

LOCUS	RESULT 11	373 bp	mrna	linear	EST 15-MAR-2007
AF433473					

DEFINITION	AA333473 S00011 Hordeum vulgare cDNA clone S0001100036E05F1, mRNA sequence.
ACCESSION	AA333473
VERSION	AA333473.1
KEYWORDS	GI:19521925
SOURCE	BSF.
	Hordeum vulgare

ORGANISM	REFERENCE
<i>Hordeum vulgare</i>	1 (bases 1 to 373)
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Hordeum.	Saren, A.-M., Tanskanen, J., Paulin, L. and Schulman, A.H.
AUTHORS	Barley EST's
TITLE	Unpublished (2002)
JOURNAL	Contact: Schulman AH
COMMENT	Institute of Biotechnology University of Helsinki P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,

FEATURES	Finland.	Location/Qualifiers
----------	----------	---------------------

	source	1..373 /organism="Hordeum vulgare" /mol_type="mRNA" /db_xref="taxon:4513" /clone="S0001100036E03F1" /dev_stage="developing seed" /clone_1fb="S00011" /note="12,15,18 days after pollination"					
ORIGIN							
	Query Match	3.1%	Score 340.4;	DB 1;	Length 373;		
	Best local similarity	99.7%;	Pred. No. 2.9e-58;				
	Matches 341;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;		
Oy	1486	CGTGGCAGACGACCTGTTCGCAGGGCGTGCCCTAATCATCTATCGGCCGACGACGCTTCTCT	1545				
Dd	28	CGAGGACGAGCACTGTTCGCAGGGCGTGCCCTAATCATCTATCGGCCGACGACGCTTCTCT	87				
Oy	1546	CACCGCCGCGGAGCCAGCGGATCGAGGTATTAATCCCAGTCTGTGGCCCCAGACC GGCG	1601				

Db 88 CACCCGCGGAGACCAAGCCGATGAGTCAATATCCGGTGTGGCCCAAGACGCGCG 147
Qy 1606 CGCCAAATTTGCTCTCCCAAGCCAGAGCTACCCAACTGAGAGCGCGCGCTTCAA 1665
Db 148 CGCCAAATTTGCTCTCCCAAGCCAGAGCTACCCAACTGAGAGCGCGCGCTTCAA 207
Qy 1666 CAGGCTGAGAGTCCGAGCATTTGACCTCATCCCGCAAGAGGGTGGAGATGACATCGA 1725
Db 208 CAGGCTGAGAGTCCGAGCATTTGACCTCATCCCGCAAGAGGGTGGAGATGACATCGA 267
Qy 1726 CTGCTGGAATTCATCGCCGACAAACAACACCGGCATGTCATCATTAACCCCAACA 1785
Db 268 CTGCTGGAATTCATCGCCGACAAACAACACCGGCATGTCATCATTAACCCCAACA 327
Qy 1786 CCGGTCGGGACGCTTACTCTAGACCATCTGTCCAGGT 1827
Db 328 CCGGTCGGGACGCTTACTCTAGACCATCTGTCCAGGT 369

RESULT 12
LOCUS CD878050 635 bp mRNA linear EST 11-JUL-2003
DEFINITION AZ04.101M17P011002 AZ04 Triticum aestivum cDNA clone AZ04101M17,
mRNA sequence.
ACCESSION CD878050
VERSION CD878050.1 GI:32561866
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
1 (bases 1 to 635)
Genoplante.
REFERENCE Genoplante, a major partnership french program in plant genomes
Unpublished (2003)
AUTHORS Contact: Genoplante
JOURNAL
COMMENT Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genome programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiolegen.fr>).

FEATURES
Source location/Qualifiers
1..635
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="rectal"
/db_xref="taxon:4565"
/clone="AZ04101M17"
/issue_type="root"
/clone_11b="AZ04"

ORIGIN
Query Match 3.1%; Score 340; DB 6; Length 635;
Best local similarity 90.8%; Pred. No. 3.8e-58;
Matches 374; Conservative 0; Mismatches 35; Indels 3; Gaps 1;
Qy 891 ACGGACAGACCAAGCCGACCGCGAGACCAACGCGCATGTCAGGCGCGCGAGAGCG 950
Db 96 ACGGACCGCGCGCGCGCAACCGCAAGACCAAGCGCGCATGTCAGGCGCGCGAGAGCG 155
Qy 951 GCGAGAGCAAGAGATGCTCGAGAGCTCCGCGCGCAACGCGCGAGCAACGCGCGCTGCGG 1010
Db 156 GCGAGAGCAAGAGCGCGCGCGCGCGCTCTGTGTAAGCGCGAGCAACGCGCGCTGCGG 215
Qy 1011 GCGGCGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1067
Db 216 GCGGCGCGCGCGCGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 275
Qy 1068 GCGGTCGGCGCGCGAGCGGCGCGAGCATGAGCATCCGGCGATACGTTACAGATCAGCG 1127

Db 276 GGTCTGCGCGCGCGAGCGGCGCGGAGACATACGATCCGGCGATACGTTCAAGATCAGCG 335
Qy 1128 CGAGGTCGAGAGAGAGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1187
Db 336 CGAGGTCGAGAGAGAGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395
Qy 1188 TGTTCGCGCGCTTCCGCAACGCGCGCTGAGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 1247
Db 396 TGTTCGCGCGCTTCCGCAACGCGCGCTGAGAGCGCGAGAGTGTGCTGCTGCTGCTGCTGCTG 455
Qy 1248 CCGGCAAGTTCAATGCTACCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1299
Db 456 CCGGCAAGTTCAATGCTACCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507

RESULT 13
LOCUS AJ485409 360 bp mRNA linear EST 24-MAY-2002
DEFINITION AJ485409 S00011 Hordeum vulgare cDNA clone S0001100055B10P1, mRNA
sequence.
ACCESSION AJ485409
VERSION AJ485409.1 GI:21201364
KEYWORDS EST.
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
1 (bases 1 to 360)
Genoplante.
REFERENCE Saren, A.-M., Tanekannen, J., Paulin, L. and Schulman, A.H.
Unpublished (2002)
AUTHORS Contact: Schulman AH
JOURNAL Institute of Biotechnology
University of Helsinki
P.O. Box 56 (Valkankari 6A), University of Helsinki FIN-00014,
Finland.

FEATURES
Source location/Qualifiers
1..360
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
/clone="S0001100055B10P1"
/dev_stage="Developing seed"
/clone_11b="S00011"
/note="12,15,18 days after pollination"

ORIGIN
Query Match 3.1%; Score 337; DB 1; Length 360;
Best local similarity 100.0%; Pred. No. 1.4e-57;
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1491 CAGAGCACTGTCGAGAGGCGTGCCTACAGCTATCGGCGAGCAGCTCTTCTCACCG 1550
Db 1 CAGAGCACTGTCGAGAGGCGTGCCTACAGCTATCGGCGAGCAGCTCTTCTCACCG 60
Qy 1551 CCGGCGGAGCCAGGCGATGATGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1610
Db 61 CCGGCGGAGCCAGGCGATGATGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Qy 1611 ACATTCGCTCCCGAGCGAGGCTACCAACTAGAGGCGCGCGCGCTTCAACAGCG 1670
Db 121 ACATTCGCTCCCGAGCGAGGCTACCAACTAGAGGCGCGCGCGCTTCAACAGCG 180
Qy 1671 TGAAGTCCGAGATTTGACCTCATCCCGCAAGAGGCTGGAGATGACATGCACTGCG 1730
Db 181 TGAAGTCCGAGATTTGACCTCATCCCGCAAGAGGCTGGAGATGACATGCACTGCG 240
Qy 1731 TGAATTCATGCGCGAGCAAGAACACACCGCATGTCATATTAACCCCAACACCGCT 1790
Db 241 TGAATTCATGCGCGAGCAAGAACACACCGCATGTCATATTAACCCCAACACCGCT 300

Qy 1791 GCGGACGGTTTACTCTTACGACCATCTGTCCAAAGT 1827
 Db 301 GCGGACGGTTTACTCTTACGACCATCTGTCCAAAGT 337
 RESULT 14
 CG033764/c 940 bp DNA linear GSS 19-AUG-2003
 LOCUS PUFYB66TB.ZM.0.6.1.0.Zea.mays.genomic.clone.ZM59K12,
 DEFINITION genomic survey sequence.
 ACCESSION CG033764
 VERSION CG033764.1 GI:33905920
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 940)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennerzen,J.
 REFERENCE Unpublished (2003)
 JOURNAL Other_GSSs: PUFYB66TD
 COMMENT Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteLaw@tigr.org
 Seq primer: TR
 Class: sheared ends.
 FEATURES
 source location/Qualifiers
 1..940
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZM59K12"
 /clone_1lb="ZM.0.6.1.0.KB"
 /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
 Cor selected genomic DNA library"
 ORIGIN
 Query Match 3.0%; Score 325.6; DB 9; Length 940;
 Best Local Similarity 70.0%; Pred. No. 3.6e-55;
 Matches 483; Conservative 0; Mismatches 199; Indels 8; Gaps 3;
 Qy 7012 CTTATTAACAAACAGCGCGGTAGACAGCATTTGTACAGGCGGTGCTTACAAAGCTA 7071
 Db 804 CTTTAAAAAACAAGTGTCTGTGCGGAGCACCTGTCAAGTGAATCTTCAATACAAAGCTG 745
 Qy 7072 TCGGCGCGACGAGTCTTCTCAACGCGCGGGAATCTCAGGCGGATGAGTCAATATCCCG 7131
 Db 744 TCAACTGATGACATCTTCTCAACGCGGTAGACATCAAGCATTTGAGTGTGCTCA 685
 Qy 7132 GTGTGCGCCAGACTGCGCGGCGCAACATCTTCTCCCGCGAGGCTATCAAAATTAC 7191
 Db 684 GTCTTGCCCA--ACGGGGCGCAACATATTGTCTCCCAACAGAGCTATCCAAATTAT 628
 Qy 7192 GAGGCGGAGCGGCAATTCACAGAGTGAAGTCCGGCACTTGCACCTCATCCCGACAG 7251
 Db 627 GAGGCGGCTGACGAGCTGCACAACTTACAAAGTGTGATTCGATCTGAGAGA 568
 Qy 7252 GGGTGGGAGATCGACATGATCCCTGGAATTCATCCGCGCAAGAAACACACCGCGCAT 7311
 Db 567 GGGTGGGAAATCGAATGACCTCTGAGTGTGACAAAGAACACACCGCAATG 508
 Qy 7312 GTCAATCAAAACCAAAACATCCGTCGCGGAGCGTTTACTCTTACGACCATCTGGCCAG 7371
 Db 507 GTCAATCAAAACCAAAACATCTTGGCGGAGTGTCTACCCCGAATTTAGCCAG 448

Qy 7372 GTTTGATCANGATCCTCTGCGCTTGATGACCGGCTGTGTAACATAGTATAT 7431
 Db 447 GCATGGCTGCTTATCTCATGACCATTTAAATGAATTAAGTTAATTAAGTCTTC 388
 Qy 7432 GAATTGGCTTTCATTCGTGTGCTGATGATGCTGTTTGTTA--TCAGTGGCGGAGCTG 7490
 Db 387 GCGGAGAC---CAATTTGTTTAAAGACAGCTTTATATGATTAAGTGGCGGAGGTA 332
 Qy 7491 GCAAGAACTCGAATATTTGATGATGCTGACGACGATTTACGCGAAACTGTTCTGGGC 7550
 Db 331 GCAAGAACTGGAATTAATGATGATGCTGATGATGATGATGATGATGATGATGATGAT 272
 Qy 7551 AGCGCCCGCTTATCCGATGAGCGCTCTTGGGACATGCCCCGCTTTGTCCATTGGA 7610
 Db 271 GACACCCCTTACGTCCTCAATGGGTGTCTTGGCCACATGCCCCGCTTTGACATAGGA 212
 Qy 7611 TCTCTGCAAGTGTGATGATGCTGATGAGGACCTGATGAGTGGGCGGTGATGAC 7670
 Db 211 TCACTATGAAAGAGTGAATGATGCTGAGGAGGACCTGATGATGATGATGATGATGAT 152
 Qy 7671 CCCACAAAGATTTAGAGAAACTAGATA 7700
 Db 151 CCCACAAAGATTTGCAAGAAACCAAGATA 122
 RESULT 15
 CC696106/c 934 bp DNA linear GSS 19-JUN-2003
 LOCUS OG10151TV.ZM.0.7.1.5.Zea.mays.genomic.clone.ZM59K12,
 DEFINITION genomic survey sequence.
 ACCESSION CC696106
 VERSION CC696106.1 GI:32100882
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 934)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Reenick,A., Fraser,C.M., Budiman,M., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nuberger,A., Robbins,D. and Lakey,N.
 REFERENCE Unpublished (2002)
 JOURNAL Other_GSSs: OG10151TH
 COMMENT Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteLaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 FEATURES
 source location/Qualifiers
 1..934
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZM59K12"
 /clone_1lb="ZM.0.7.1.5.KB"
 /note="Vector: pCR4-TOPO; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"
 ORIGIN
 Query Match 2.9%; Score 316.6; DB 9; Length 934;
 Best Local Similarity 71.4%; Pred. No. 2.4e-53;
 Matches 445; Conservative 0; Mismatches 174; Indels 4; Gaps 2;
 Qy 1467 AAACAAAAAACAAGCGCGGTGCAAGCACTGTGCAAGGCGGTGCTTACATGCTAT 1526
 Db 620 AAACAAAAAACAAGTGTCTGTGCGGAGCACCTGTGAGGACCTTCAATACAGCTGT 561

Result No.	Score	Match	Length	DB	ID	Description
1	20.4	85.0	999	4	BG334930	BG334930 602461255
2	19.8	82.5	563	1	AJ807947	AJ807947 AJ807947
3	19.4	80.8	723	5	BU612123	BU612123 UI-M-EWO-
4	19.4	80.8	946	5	BUS05136	BUS05136 AGENCOURT
5	19.2	80.0	198	1	AA359060	AA359060 EST88008
6	19.2	80.0	224	1	AA359061	AA359061 EST88009
7	19.2	80.0	279	7	CN495808	CN495808 Mdlw2020P
8	19.2	80.0	478	8	AO880175	AO880175 HS 5037 B
9	18.8	78.3	391	2	BE715690	BE715690 CM2-HT075
10	18.8	78.3	444	5	BD015926	BD015926 UI-H-DPI-
11	18.8	78.3	541	2	BE706168	BE706168 RCI-HT022
12	18.8	78.3	579	8	B05889	B05889 CSRL-72911-
13	18.8	78.3	654	4	BG334045	BG334045 602460411
14	18.8	78.3	926	5	BK371803	BK371803 BX371803
15	18.8	78.3	940	5	BK371804	BK371804 BX371804
16	18.8	78.3	989	5	B0068921	B0068921 AGENCOURT
17	18.8	78.3	1004	5	B0069095	B0069095 AGENCOURT
18	18.8	78.3	1190	8	CC242470	CC242470 CH261-11F
19	18.8	78.3	1403	3	HSM805111	BLX333333 Homo sapi
20	18.8	78.3	8664	3	BSM080880	BSM080880
21	18.4	76.7	532	4	BG996584	BG996584 QV4-HT122
22	18.4	76.7	552	8	B2266969	B2266969 CH230-254
23	18.4	76.7	571	8	AQ465392	AQ465392 HS 5119 A
24	18.2	75.8	240	2	BE932400	BE932400 QV3-HT063
25	18.2	75.8	442	2	BE509060	BE509060 dc16e06-Y
26	18.2	75.8	451	8	B64254	B64254 CIT-HSP-202
27	18.2	75.8	463	7	CY359206	CY359206 PMO-CT041
28	18.2	75.8	530	4	BM728080	BM728080 UI-E-EUO-
29	18.2	75.8	596	1	AL791979	AL791979 AT-791979
30	18.2	75.8	600	7	CO358000	CO358000 DR AT9E OR
31	18.2	75.8	630	1	AL865290	AL865290 AL862250
32	18.2	75.8	649	1	AL859034	AL859034 AL859034
33	18.2	75.8	654	1	AL887699	AL887699 AL887699
34	18.2	75.8	658	1	AL864891	AL864891 AL864891
35	18.2	75.8	661	1	AL875887	AL875887 AL875887
36	18.2	75.8	665	1	AL861073	AL861073 AL861073
37	18.2	75.8	682	4	B1186061	B1186061 UNL-P-FN-
38	18.2	75.8	688	7	CR433911	CR433911 CR433911
39	18.2	75.8	726	1	BK776131	BK776131 BK776131
40	18.2	75.8	727	5	BK731133	BK731133 BK731133
41	18.2	75.8	750	4	B1183205	B1183205 UNL-P-FN-
42	18.2	75.8	751	5	BK750164	BK750164 BK750164
43	18.2	75.8	762	7	CR433912	CR433912 CR433912
44	18.2	75.8	792	7	CK127026	CK127026 AGENCOURT
45	18.2	75.8	826	5	BK781066	BK781066 BK781066

RESULT.1	LOCUS	DEFINITION
BG334930/c	BG334930	999 bp mRNA linear EST 27-FEB-2001
	602461255F1	NIH MGC 20 Homo sapiens cDNA clone IMAGE:4578228 5',

ACCESSION mRNA sequence.
 BG334930
 VERSION BG334930.1 GI:13141368
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 999)
 NIH-MGC http://mgi.nci.nih.gov/.

REFERENCE
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DFP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Inceye Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN, at: http://image.llnl.gov
 Plate: LICM1293 row: m column: 13
 High quality sequence stop: 520.
 Location/Qualifiers
 1..999
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4578228"
 /tissue_type="melanocytic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH MGC 20"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
 Query Match 85.0%; Score 20.4; DB 4; Length 999;
 Best Local Similarity 95.5%; Pred. No. 1.1e+02;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 TCTGGATCTCAGTCATCA 22
 Db 567 TCTGGATCTCAGTCATCA 546

RESULT 2
 AJ807947 563 bp mRNA linear EST 11-MGC-2004
 LOCUS AJ807947
 DEFINITION Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018_6_08 d24, mRNA sequence.
 ACCESSION AJ807947
 VERSION AJ807947.1 GI:51123275
 KEYWORDS EST.
 SOURCE Antirrhinum majus (snapdragon)
 ORGANISM Antirrhinum majus

REFERENCE
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Plantales; Plantaginaceae; Antirrhineae; Antirrhinum.
 1 (bases 1 to 563)
 Zenghe, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.
 Antirrhinum EST Collection
 Unpublished (2003)
 Contact: Schwarz-Sommer Z
 Molekulare Pflanzen-genetik
 MPI fuer Zuechtungs-forschung
 Carl-von-Linne Weg 10, D-50829, Germany.

FEATURES
 source Location/Qualifiers
 1..563
 /organism="Antirrhinum majus"
 /mol_type="mRNA"
 /db_xref="taxon:4151"
 /clone="018_6_08 d24"
 /tissue_type="whole plant"
 /clone_1lb="Antirrhinum majus whole plant"

ORIGIN
 Query Match 82.5%; Score 19.8; DB 1; Length 563;
 Best Local Similarity 91.3%; Pred. No. 1.9e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 2 CTGGATCTCAGTCATCA 24
 Db 233 CTGGATCTCAGTCATCA 255

RESULT 3
 BU612123 723 bp mRNA linear EST 20-FEB-2003
 LOCUS BU612123
 DEFINITION UT-M-EMO-cax-k-01-0-UT.r1 NIH BMAP EMO Mus musculus CDNA clone UT-M-EMO-cax-k-01-0-UT 5', mRNA sequence.
 ACCESSION BU612123
 VERSION BU612123.1 GI:23278338
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 723)
 NIH-MGC http://mgi.nci.nih.gov/.

REFERENCE
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..723
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="UT-M-EMO-cax-k-01-0-UT"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1lb="NIH BMAP EMO"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGTGGA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 80.8% Score 19.4; DB 5; Length 723;
 Best Local Similarity 95.2%; Pred. No. 36+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGGATCTCAGATCCACCA 24
 |||||
 DB 475 GGGATCTCAGATCCACCA 495

RESULT 4
 BUS05136 946 bp mRNA linear EST 12-SEP-2002
 LOCUS AGENCOURT 10013187 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6493080
 DEFINITION 5', mRNA sequence.

ACCESSION BUS05136
 VERSION BUS05136.1 GI:22811369
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 AUTHORS NIH-MGC http://mgi.mgi.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strusberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: The Cepho Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL site:
 http://image.llnl.gov
 Plate: L14M14047 row: c column: 01
 High quality sequence stop: 628.

FEATURES

source 1..946
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6493080"
 /ciseue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /clone_11b="NIH_MGC_94"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 80.8% Score 19.4; DB 5; Length 946;
 Best Local Similarity 95.2%; Pred. No. 3.1e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GGGATCTCAGATCCACCA 24
 |||||
 DB 689 GGGATCTCAGATCCACCA 709

RESULT 5
 AA359060/c 198 bp mRNA linear EST 21-APR-1997
 LOCUS EST68008 Fetal lung II Homo sapiens cDNA 5' end, mRNA sequence.

ACCESSION AA359060
 VERSION AA359060.1 GI:2011377
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 198)
 Adams M.D., Kerlavage A.R., Fleischmann R.D., Fuldner R.A.,
 Bult C.J., Lee N.H., Kirkness E.F., Weinstock K.G., Gocayne J.D.,
 White O., Sutton G., Blake J.A., Brandon R.C., Man Wal C.,
 Clayton R.A., Cline T.R., Cotton M.D., Earle-Hughes J., Fine L.D.,
 Fitzgerald L.M., Fitzhugh W.M., Fitchman J.L., Geoghegan N.S.,
 Glodex J.M., Gnehm C.L., Hanna M.C., Hedblom E., Hinkle P.S., Jr.,
 Kelley J.M., Kelley J.C., Liu L.-I., Marmaros S.M., Merrick J.M.,
 Moreno-Palanges R.F., McDonald L.A., Nguyen D.T., Pellegrino S.M.,
 Phillips C.A., Ryder S.E., Scott J.L., Soudk D.M., Shirley R.,
 Small K.V., Spriggs T.A., Utterback T.R., Weidman J.F., Li Y.,
 Bednarek D.P., Cao L., Cepeda M.A., Coleman T.A., Collins E.J.,
 Dimke D., Feng D.-F., Ferris A., Fischer C., Hastings G.A.,
 He W.M., Hu J.S., Greene J.M., Gruber J., Hudson P., Kim A.K.,
 Kozak D.L., Kunach C., Hungjun J., Li H., Meisner P.S., Olsen H.,
 Raymond L., Wei X.F., Wang J., Xu C., Yu G.H., Ruden S.W.,
 Dillon P.J., Fannon M.R., Rosen C.A., Haseltine W.A., Fields C.,
 Fraser C.M., and Venter J.C.
 Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 96026280
 7566098

TITLE

JOURNAL
 MEDLINE
 PUBMED

COMMENT

Other ESTs: TRC76983
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org

For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (http://www.tigr.org/cdb/hgi/hgi.html)
 Seq primer: M13 Reverse.

FEATURES

source 1..198
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):163152"
 /db_xref="taxon:9606"
 /sex="male"
 /dev stage="fetus, 19 wks"
 /clone_11b="Fetal lung II"
 /note="Organ: lung; Vector: pBlueScript SK-; Site_1:
 EcoRI; Site_2: XhoI"

ORIGIN

Query Match 80.0% Score 19.2; DB 1; Length 198;
 Best Local Similarity 87.5%; Pred. No. 3.2e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCTGGATCTCAGATCCACCA 24
 |||||
 DB 162 TCTGGATCTCAGATCCACCA 139

RESULT 6
 AA359061/c 224 bp mRNA linear EST 21-APR-1997
 LOCUS EST68009 Fetal lung II Homo sapiens cDNA 5' end, mRNA sequence.

ACCESSION AA359061
 VERSION AA359061.1 GI:2011378
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 224)
 Adams M.D., Kerlavage A.R., Fleischmann R.D., Fuldner R.A.,
 Bult C.J., Lee N.H., Kirkness E.F., Weinstock K.G., Gocayne J.D.,
 White O., Sutton G., Blake J.A., Brandon R.C., Man Wal C.,
 Clayton R.A., Cline T.R., Cotton M.D., Earle-Hughes J., Fine L.D.,

Fitzgerald, J.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodex, A., Gnehm, C.L., Hanna, M.C., Hedblom, B., Hinkle, P.S., Jr., Kallek, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palanges, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.B., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Korak, D.L., Kunach, C., Hungjun, J., Li, H., Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

JOURNAL
MEDLINE
96026280
7566098
Other ESTs: THCT6983
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arterlav@igr.org
For clone availability, additional sequence and expression information available to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/cdb/hgt/hgt.html>)
Seq primer: M13 Reverse.

FEATURES
source
location/Qualifiers
1..224
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):163153"
/db_xref="taxon:9606"
/sex="male"
/dev_stage="fetus, 19 wks"
/clone_lib="Petal lung II"
/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 80.0%; Score 19.2; DB 1; Length 224;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCTGGGATCTCAGTCATCCACA 24
|||||
Db 162 TCTGGGCTTTCAGTCATCCATCA 139

RESULT 7
CN495808/c 279 bp mRNA linear EST 24-MAY-2004
LOCUS Mdfw2020p20.y1 Mdfw Malus x domestica cDNA clone Mdfw2020p20.5'
DEFINITION similar to TR:Q92P29 Q92P29 PUTATIVE GRI1-LIKE PROTEIN.; mRNA
sequence.
CN495808
ACCESSION CN495808.1 GI:46597534
VERSION
KEYWORDS EST.
SOURCE Malus x domestica (cultivated apple)
ORGANISM Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustroids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 279)
Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A., Alwinckle, H., Malinov, M., Carroll, N., Goldsbrough, P., Oryis, K., Clifton, S., Pape, D., Maria, M., Hillier, L., Martin, J., Wylie, T., Darte, M., Theising, B., Bowers, Y., Gibbons, M., Ritzer, E., Ronko, I., Tsagaris, V., Kennedy, S., Waterson, R., and Wilson, R.
Apple Functional Genomics grant - NSF 0321702

JOURNAL
COMMENT
Unpublished (2004)
Contact: Schuyler S. Korban
Apple Functional Genomics grant - NSF 0321702
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: esat@watson.wustl.edu
Library materials provided by: Schuyler S. Korban Library
constructed by: A. Hernandez / K. Gasic Library sequenced by:
Washington University Genome Sequencing Center
WashU EST name: aa180h10.y1
High quality sequence stop: 255.

FEATURES
source
location/Qualifiers
1..279
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="Mdfw2020p20"
/lab_host="DH10B ampicillin resistant"
/clone_lib="Mdfw"
/note="Vector: DH10B ampicillin resistant; Site_1: NotI; Site_2: EcoRI; Total RNA was extracted separately from each stage (bud, balloon, open and after pollination) using the 'pine tree' method. Poly(A)+mRNA was isolated twice from total RNA from each stage using the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse transcribed into double stranded cDNA using a modified oligo(dT) primer with an identifying tag sequence (see table below). cDNAs from different stages were pooled in equal amounts before adaptor ligation. Tag identification when sequencing from 5' end: Stage 1 (bud) insert 18(A)TGCAG; Stage 2 (balloon) insert 18(A)TGCAG; Stage 3 (open) insert 18(A)TGCCT; Stage 4 (after pollination) insert 18(A)TGCAG; Tag identification when sequencing from 3' end: Stage 1 (bud) TCCGAl8(T) insert; Stage 2 (balloon) TCCGAl8(T) insert; Stage 3 (open) ACCGAl8(T) insert; Stage 4 (after pollination) ACCGAl8(T) insert. Double stranded cDNAs were size selected (more than 450 bp), adapted with EcoRI adapters at both ends and then digested with NotI. The cDNAs were then directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector (Stratagene). Identification of adaptors and tags in 5'-end sequenced clones: <Vector>...TAGCTT<End Vector><Start EcoRI adaptor>GATTCGAATTCATTGTGTGGG<End EcoRI adaptor><Start Insert>...AAAAAAAAAAAAAAAA<End Insert><Start Tag>TGCAG<End Tag><Start NotI site>Vector>GCCGCCGCCACCGCG... The total number of white colony forming units (cfu) in the primary library before amplification was 1.1x10⁶ cfu (colony forming units). The background of empty clones was less than 1%. Inserts ranged from 0.5kb to 3 kb, as determined by PCR. Purified plasmid DNA from the primary library was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites flanking the cloned cDNA inserts. The purified PCR products, representing the entire cloned cDNA population, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30C. Unhybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B cells to generate the normalized library. The total number of clones with insert was 9x10⁶ cfu. Background of empty clones was less than 1%.

ORIGIN
Query Match 80.0%; Score 19.2; DB 7; Length 279;
Best Local Similarity 87.5%; Pred. No. 3.3e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCTGGGATCTCAGTCATCCACA 24
|||||
|||||

Db 197 TCTGTGATCTGAGTCATCCACAGA 174

RESULT 8
LOCUS A0880175 478 bp DNA 1linear GSS 09-NOV-1999
DEFINITION HS 5037 B1.D11.T7 RPCR-11 Human Male BAC Library Homo sapiens genomic clone Plate=8805 Col=21 Row=H, genomic survey sequence.
ACCESSION A0880175
VERSION A0880175.1 GI:6311642
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 478)
Mahalas,G.G., Wallace,J.C., Sutch,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahalas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCR-11. For BAC library availability, please contact Plesier de Jong (plesier@edisons.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (<http://www.resgen.com>). BAC end web server: <http://www.htbc.washington.edu>
Plate: 8805 row: H column: 21
Seq primer: T7
Class: BAC ends
High quality sequence stop: 478.
Location/Qualifiers
1..478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=8805 Col=21 Row=H"
/sex="male"
/clone_1lb="RPCR-11 Human Male BAC Library"
/note="Vector: PBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the PBAC3.6 vector at EcoRI sites"

FEATURES
source

ORIGIN
1..478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=8805 Col=21 Row=H"
/sex="male"
/clone_1lb="RPCR-11 Human Male BAC Library"
/note="Vector: PBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the PBAC3.6 vector at EcoRI sites"

Query Match 80.0%; Score 19.2; DB 8; Length 478;
Best Local Similarity 87.5%; Pred. No. 3.6e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TCTGGATCTCAGTCATCCACACA 24
|||||
313 TCTGGAGCTCATTCATCTCTACA 336
|||||

Db 313 TCTGGAGCTCATTCATCTCTACA 336
|||||

RESULT 9
LOCUS BE715690/c 391 bp mRNA 1linear EST 12-SEP-2000
DEFINITION CM2-HT0750-040700-250-cl1 HT0750 Homo sapiens CDNA, mRNA sequence.
ACCESSION BE715690
VERSION BE715690.1 GI:10103955
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 391)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongseneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.U.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.U.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704932
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?rl=kr2-cm2-HT0750-040700-250-cl1&cl=2000-07-04&f4=1>)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 391.
Location/Qualifiers
1..391
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1lb="HT0750"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORFESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source

ORIGIN
1..391
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1lb="HT0750"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORFESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Query Match 78.3%; Score 18.8; DB 2; Length 391;
Best Local Similarity 90.9%; Pred. No. 5.4e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTGGATCTCAGTCATCCAC 23
|||||
228 CTGGGTTCTCAGTCATCCAGC 207
|||||

Db 228 CTGGGTTCTCAGTCATCCAGC 207
|||||

RESULT 10
LOCUS B0015926 444 bp mRNA 1linear EST 17-JUN-2002
DEFINITION UT-H-DT1-gvz-d-05-0-UT.g1 NCI_CGAP_DTI Homo sapiens CDNA clone IMAGE:5886292 3', mRNA sequence.
ACCESSION B0015926
VERSION B0015926.1 GI:19751203
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 444)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..444

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5886292"

/tissue_type="Metastatic Chondrosarcoma"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_id="NCI_CGAP DT1"

/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: Ecor I; Site 2: Not I;

NCI CGAP DT1 is a normalized CDNA library containing the

following tissue(s): Metastatic Chondrosarcoma in Lung.

The library was constructed according to Bonaldo, Lennon

and Soares, Genome Research, 6:791-806, 1996. First strand

CDNA synthesis was primed with an oligo-dT primer

containing a Not I site. Double stranded cDNA was ligated

to an Ecor I adaptor, digested with Not I, and cloned

directionally into pT7T3-Pac vector. The oligonucleotide

used to prime the synthesis of first-strand cDNA contains

a library tag sequence that is located between the Not I

site and the (dT)18 tail. The sequence tag for this

library is AACTGTCGC.

TAG_TISSUE=Lung metastatic chondrosarcoma

TAG_LIB=UI-H-DT1

TAG_SEQ=AACTGTCGC"

ORIGIN

Query Match

Best Local Similarity 78.3%; Score 18.8; DB 5; Length 444;

Best Local Similarity 90.9%; Pred. No. 5.5e+02;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Dy 2 CTGGGATCTCAGTCATCAAC 23

Db 400 CTGGGATCTCAGTCATCAAC 421

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

FEATURES

source

Location/Qualifiers

1..541

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_id="HT0229"

/note="Organ: head, neck; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: astampson@ludwig.org.br
 This sequence was derived from the PAPSP/ICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scrlp/getshtml2.pl?l=et2-RC1-HT0229-160>)
 600-112-6074c3-2000-06-16ct4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 541.

FEATURES

source

Location/Qualifiers

1..541

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_id="HT0229"

/note="Organ: head, neck; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: astampson@ludwig.org.br
 This sequence was derived from the PAPSP/ICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scrlp/getshtml2.pl?l=et2-RC1-HT0229-160>)
 600-112-6074c3-2000-06-16ct4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 541.

FEATURES

source

Location/Qualifiers

1..579

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_id="HT0229"

/note="Organ: head, neck; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: astampson@ludwig.org.br
 This sequence was derived from the PAPSP/ICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scrlp/getshtml2.pl?l=et2-RC1-HT0229-160>)
 600-112-6074c3-2000-06-16ct4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 541.

FEATURES

source

Location/Qualifiers

1..579

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_id="HT0229"

/note="Organ: head, neck; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

low stringency conditions."

/clone="CSRL-72g11"
/sex="female"
/cell type="chimeric hamster somatic cell hybrid"
/clone.lib="CSRL flow sorted Chromosome 11 specific
cosmid"
/note="Vector: sCos-1; Human Chromosome 11 specific cosmid
library prepared from flow sorted human Chromosome 11
derived from Chinese Hamster Ovary (CHO) monochromosomal
somatic cell hybrid, J1"

ORIGIN

Query Match 78.3%; Score 18.8; DB 8; Length 579;
Best Local Similarity 90.9%; Pred. No. 5.7e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGGGATCTCAGTCGATCCAC 23
DB 239 CTGGGTTCTCAGTCGATCCAGC 218

RESULT 13

LOCUS BG334045 654 bp mRNA linear EST 27-FEB-2001
DEFINITION 602460411P1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4577258 5',
mRNA sequence.

ACCESSION BG334045
VERSION BG334045.1 GI:13140483
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 654)
NIH-MGC <http://mgs.nci.nih.gov/>.
Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DTP

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LCM1291 row: e column: 03
High quality sequence stop: 652.
Location/Qualifiers

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4577258"
/cfeature="IMAGE:4577258"
/cfeature="IMAGE:4577258"
/lab_host="DH10B (phage-resistant)"
/clone.lib="NIH MGC 20"
/note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 78.3%; Score 18.8; DB 4; Length 654;
Best Local Similarity 90.9%; Pred. No. 5.8e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGGGATCTCAGTCGATCCAC 23
DB 560 CTGGGTTCTCAGTCGATCCAGC 539

RESULT 14

BX371803

LOCUS BX371803 926 bp mRNA linear EST 28-APR-2004
DEFINITION BX371803 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS001061Y108 3-PRIME, mRNA sequence.

ACCESSION BX371803
VERSION BX371803.2 GI:46831284
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 926)
Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-G1190(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
5982.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0BA1024ZFO7_CS02262_1&c=5982.f

FEATURES

source

Location/Qualifiers
1..926
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS001061Y108"
/cfeature="CS001061Y108"
/cfeature="PLACENTA COT 25-NORMALIZED"
/clone.lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-G1190(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 78.3%; Score 18.8; DB 5; Length 926;
Best Local Similarity 90.9%; Pred. No. 6.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTGGGATCTCAGTCGATCCAC 23
DB 503 CTGGGTTCTCAGTCGATCCAGC 524

RESULT 15

BX371804

LOCUS BX371804 940 bp mRNA linear EST 28-APR-2004
DEFINITION BX371804 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS001061Y108 3-PRIME, mRNA sequence.

ACCESSION BX371804
VERSION BX371804.1 GI:30456058
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 940)
Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by life technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 5982.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0BA1024ZF07_CS02262_2&c=5982.f

FEATURES

Source

Location/Qualifiers
 1. 940
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1061Y108"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 78.3%; Score 18.8; DB 5; Length 940;
 Best Local Similarity 90.9%; Pred. No. 6.1e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTGGGATCTCAGTGCATCCAC 23
 |||||
 Db 507 CTGGGTCTCAGTGCATCCAGC 528

Search completed: October 15, 2005, 11:53:26
 Job time : 73.582 secs

[illegible]

Db	1981	GCTGTTCTGGGCAAGCCGCCGCTTCATCCCAATGGGAATGTTTGGGCACATCACCCCTCTG	2040
Oy	2041	GCTGTCATAGGGTCTCTGTCCAAAGTCATGGATAGTGTCTGGATGGCGGCTTGGATGGGT	2100
Db	2041	GCTGTCATAGGGTCTCTGTCCAAAGTCATGGATAGTGTCTGGATGGCGGCTTGGATGGGT	2100
Oy	2101	AGCGGTGATGACCCCGAAGAAGTCTTACAGAAAATTAAGGTACTTAAATCTCTATATCA	2160
Db	2101	AGCGGTGATGACCCCGAAGAAGTCTTACAGAAAATTAAGGTACTTAAATCTCTATATCA	2160
Oy	2161	TTCTTTTCAAAATGCTACTAAGGTGATTAAATAGTACTCTGTACAATATATTTGCTAAAT	2220
Db	2161	TTCTTTTCAAAATGCTACTAAGGTGATTAAATAGTACTCTGTACAATATATTTGCTAAAT	2220
Oy	2221	TTGTATCTGACATTTTGTGTAGATCTCTACATCAATTAAGAAATTAAGCTCAATGTCTCGA	2280
Db	2221	TTGTATCTGACATTTTGTGTAGATCTCTACATCAATTAAGAAATTAAGCTCAATGTCTCGA	2280
Oy	2281	CAGAACCCAGACACTTCAATTAGAGTCAGTCTTTGGTATTTACCTTCGTTTCAAGAAATTA	2340
Db	2281	CAGAACCCAGACACTTCAATTAGAGTCAGTCTTTGGTATTTACCTTCGTTTCAAGAAATTA	2340
Oy	2341	GTCCTTGGTATTTATCTCCCTCCCTGTCCATTTTGTCCGGTCCCATAGTTGTAGACAGCC	2400
Db	2341	GTCCTTGGTATTTATCTCCCTCCCTGTCCATTTTGTCCGGTCCCATAGTTGTAGACAGCC	2400
Oy	2401	CACGTGCATGTCAAGTGAACCGTTTTTTCACATTAAGTTTGAAGTCAAAAGTCAGACACAT	2460
Db	2401	CACGTGCATGTCAAGTGAACCGTTTTTTCACATTAAGTTTGAAGTCAAAAGTCAGACACAT	2460
Oy	2461	ACACTTGTAGTTATTTTACCTTTGTTGCTTTGATCCGATATAAATTAATAAATTAACAATA	2520
Db	2461	ACACTTGTAGTTATTTTACCTTTGTTGCTTTGATCCGATATAAATTAATAAATTAACAATA	2520
Oy	2521	CTGAACCTACTGTGAATATAATTAACAACCTGTTCTTCAAGATATATACATGTTGGCATAGGG	2580
Db	2521	CTGAACCTACTGTGAATATAATTAACAACCTGTTCTTCAAGATATATACATGTTGGCATAGGG	2580
Oy	2581	CATGCGCATATCTTTTGGGTCAAGATGACAGTATAGTTGGAACCTTTTGAATAAATAGAT	2640
Db	2581	CATGCGCATATCTTTTGGGTCAAGATGACAGTATAGTTGGAACCTTTTGAATAAATAGAT	2640
Oy	2641	ACATTGTATCTATGAGTATACCAATTTTATTAAGAAATTTCAATTTTGATATCTTGATGT	2700
Db	2641	ACATTGTATCTATGAGTATACCAATTTTATTAAGAAATTTCAATTTTGATATCTTGATGT	2700
Oy	2701	ATTGTTCTCTTGTGATTCACACGATTTACTGTGTGTTTTGTACTATCAAAATTTGTTGAG	2760
Db	2701	ATTGTTCTCTTGTGATTCACACGATTTACTGTGTGTTTTGTACTATCAAAATTTGTTGAG	2760
Oy	2761	GCAGGTCTTCCGACGATCTTGAAGAACAAGGAAGTTCTTTAAGGCGCATTAATGGT	2820
Db	2761	GCAGGTCTTCCGACGATCTTGAAGAACAAGGAAGTTCTTTAAGGCGCATTAATGGT	2820
Oy	2821	CTGCTAAAGGATCATCAGAGATATGCTACAAACAATAAAGGAAAACAATATACATTACA	2880
Db	2821	CTGCTAAAGGATCATCAGAGATATGCTACAAACAATAAAGGAAAACAATATACATTACA	2880
Oy	2881	TGTCTCTCAAGCCAGAGATCAATGTTTGTCAATGTAACCTTAATTTGTGAAGTAAAA	2940
Db	2881	TGTCTCTCAAGCCAGAGATCAATGTTTGTCAATGTAACCTTAATTTGTGAAGTAAAA	2940
Oy	2941	AAATCTTAAGGAGTCAATATATCAATTAATTAATAGATTAATCTGGGACCGAA	3000
Db	2941	AAATCTTAAGGAGTCAATATATCAATTAATTAATAGATTAATCTGGGACCGAA	3000
Oy	3001	ATGCATCCAAATATTAATCTTCAAAATTCAAATTTACATTTCTCCGACATATTTT	3060
Db	3001	ATGCATCCAAATATTAATCTTCAAAATTCAAATTTACATTTCTCCGACATATTTT	3060
Oy	3061	TGAAGATCATGTATTTTAAGATATATGACAGAGCTTAAGTTATGCTACGACTTAATCAT	3120
Db	3061	TGAAGATCATGTATTTTAAGATATATGACAGAGCTTAAGTTATGCTACGACTTAATCAT	3120

D	3061	TGAAGATGCATGTATTTTAAGAAATAATGACAGAGCTAAAGTATGTCAAGACTAATCAT	3120
Q	3121	CTGGATATCCTTTTGCCATCTTTTGTATATCTGTGGAAATGTAAATGTCAATCATATTT	3180
D	3121	CTGATATATCCTTTTGCCATCTTTTGTATATCTGTGGAAATGTAAATGTCAATCATATTT	3180
Q	3181	ACAGAAATATCATAGCTAGTTCTAGAAAGATTGATTAATTTTCTGTATACCATGAATCCTC	3240
D	3181	ACAAGAAATATCATAGCTAGTTCTAGAAAGATTGATTAATTTTCTGTATACCATGAATCCTC	3240
Q	3241	GTATTAATCTTCATGTAAACAGGTGAACCTGAACCTTACATCTTTTGGAGAAATAGACGA	3300
D	3241	GTATTAATCTTCATGTAAACAGGTGAACCTGAACCTTACATCTTTTGGAGAAATAGACGA	3300
Q	3301	TGAACATGATTTTGTCTGCGAAGCTCCGAAAAAGAAATCAAGTATCTTATGCCAGGTAG	3360
D	3301	TGAACATGATTTTGTCTGCGAAGCTCCGAAAAAGAAATCAAGTATCTTATGCCAGGTAG	3360
Q	3361	GAATCCATGTGTGATTTTGTGCTGTATATGAAGTCTTATCAATTTCCGAGATGACTATA	3420
D	3361	GAATCCATGTGTGATTTTGTGCTGTATATGAAGTCTTATCAATTTCCGAGATGACTATA	3420
Q	3421	CATATTAATGATTAACCATATATATGTGCGAAATTTGTATTAACAGGTATGAATATCTGTG	3480
D	3421	CATATTAATGATTAACCATATATATGTGCGAAATTTGTATTAACAGGTATGAATATCTGTG	3480
Q	3481	AAGACTTTTAAACACAAATATCTGTGAGAGCTAGATATCATGACTTCTCTGTGTTTC	3540
D	3481	AAGACTTTTAAACACAAATATCTGTGAGAGCTAGATATCATGACTTCTCTGTGTTTC	3540
Q	3541	TTGACCTGATGTCTTGTGTCACATGTTGTGTCTCTCACAAAAAATGCAAGCACATGT	3600
D	3541	TTGACCTGATGTCTTGTGTCACATGTTGTGTCTCTCACAAAAAATGCAAGCACATGT	3600
Q	3601	TCAATATATGTATATATATATATATTTAGACCTTAAATTAATGTGTTCAATTTAGATAT	3660
D	3601	TCAATATATGTATATATATATATATTTAGACCTTAAATTAATGTGTTCTATTTTAGATAT	3660
Q	3661	TTTTGTAGTCCAACTTATATATTTGTGTGCTATTTCTCAAAAACAAACTTATATATGTGTG	3720
D	3661	TTTTGTAGTCCAACTTATATATTTGTGTGCTATTTCTCAAAAACAAACTTATATATGTGTG	3720
Q	3721	CCTCTCAAAATGTAGGAGGTGTTCTTGGAAATGCGAAATCGGGTCCGACTTACTTTTGTCTG	3780
D	3721	CCTCTCAAAATGTAGGAGGTGTTCTTGGAAATGCGAAATCGGGTCCGACTTACTTTTGTCTG	3780
Q	3781	TGTTCCATCTCTCTTCAAGAATGTGTCTCGAAGAGTCAAAATGATTCGTCAAGAAACAA	3840
D	3781	TGTTCCATCTCTCTTCAAGATGTGTCTCGAAGAGTCAAAATGATTCGTCAAGAAACAA	3840
Q	3841	GAAAGAAATTCGAGCGATGTATCTAGTTGTATATCTGACTGAAGCTGTAAATCATTTCC	3900
D	3841	GAAAGAAATTCGAGCGATGTATCTAGTTGTATATCTGACTGAAGCTGTAAATCATTTCC	3900
Q	3901	CAGTATCCCATCTATATCTTTCAATAAATAGAACTTTAGTCTCTATGAATAAGAGT	3960
D	3901	CAGTATCCCATCTATATCTTTCAATAAATAGAACTTTAGTCTCTATGAATAAGAGT	3960
Q	3961	CAACATCTCTCTGAAATATGTTCTGTGTGTGTGTGCTCGAGCAAAACATAGGAATGTAT	4020
D	3961	CAACATCTCTCTGAAATATGTTCTGTGTGTGTGTGCTCGAGCAAAACATAGGAATGTAT	4020
Q	4021	GTTATGTAAGTTACATTTGGCGTGCAGAGATCTTTGAAGTTTTTTTTTTTTTGGGGGGGG	4080
D	4021	GTTATGTAAGTTACATTTGGCGTGCAGAGATCTTTGAAGTTTTTTTTTTTTTGGGGGGGG	4080
Q	4081	GGGGGGGGGGGTCTTGATATTTACTTTAAAGTACAGCTTCTCAAGTTATGTCAAAACA	4140
D	4081	GGGGGGGGGGGTCTTGATATTTACTTTAAAGTACAGCTTCTCAAGTTATGTCAAAACA	4140
Q	4141	CTTTGTAAACAATTGTAGATTTGGTATCATGATATGGAATTAATTAATCATGATCTTGGT	4200
D	4141	CTTTGTAAACAATTGTAGATTTGGTATCATGATATGGAATTAATTAATCATGATCTTGGT	4200

4201 AAGCAAAACCTACCTATGTTAGGCTCACTAAGGAGGCTTGGTTGAGAGAGAGA 4260
4201 AAGCAAAACCTACCTATGTTAGGCTCACTAAGGAGGCTTGGTTGAGAGAGAGA 4260
4261 GGATCAGTTGATGATATCCCAATCATCGAAGTAAATCATGTGTTGTTGCTACCACTTT 4320
4261 GGATCAGTTGATGATATCCCAATCATCGAAGTAAATCATGTGTTGTTGCTACCACTTT 4320
4321 CTACAAATCTTACTGACTGAGGCTTGAAGTACTGATCAACACCTGACACACATATT 4380
4321 CTACAAATCTTACTGACTGAGGCTTGAAGTACTGATCAACACCTGACACACATATT 4380
4381 CTCTGTGCAAAATCGGCAACCAAGATTACATCTCAAGCTGACCAACCAAAATTG 4440
4381 CTCTGTGCAAAATCGGCAACCAAGATTACATCTCAAGCTGACCAACCAAAATTG 4440
4441 AAGAGAGAAACCTCACAAAAGACCTTGAAGTCCCCCAAAATGATGTTAGGCGCG 4500
4441 AAGAGAGAAACCTCACAAAAGACCTTGAAGTCCCCCAAAATGATGTTAGGCGCG 4500
4501 TCGAGGCGGAGTGTCAACCAATGCGGACCAACCACTCCAGGGGAGACAGTAC 4560
4501 TCGAGGCGGAGTGTCAACCAATGCGGACCAACCACTCCAGGGGAGACAGTAC 4560
4561 CGATTAAGTAAATCCCAAAATCTTAAATTTGTAACCAAAATTTAAAAAGAGACA 4620
4561 CGATTAAGTAAATCCCAAAATCTTAAATTTGTAACCAAAATTTAAAAAGAGACA 4620
4621 ATTTTGAATTTGTAACCAATTTTAAACGGATATCTGAAACATTTTCAAAATTTG 4680
4621 ATTTTGAATTTGTAACCAATTTTAAACGGATATCTGAAACATTTTCAAAATTTG 4680
4681 GATCAAAATTTTAAACGACCTCTTCAAAATTTGAGCAATTTTAAATTTAAAAA 4740
4681 GATCAAAATTTTAAACGACCTCTTCAAAATTTGAGCAATTTTAAATTTAAAAA 4740
4741 GTTCAACAATTTGAACTTTTAAATTTAGCGAACAATTTGAAATTTCTAAATTTT 4800
4741 GTTCAACAATTTGAACTTTTAAATTTAGCGAACAATTTGAAATTTCTAAATTTT 4800
4801 CGAATTTGGAACATTTTCTAATTTCTGAACAAAATTTGAAATTTGAAATTTG 4860
4801 CGAATTTGGAACATTTTCTAATTTCTGAACAAAATTTGAAATTTGAAATTTG 4860
4861 ATTAATTTTGAAGATTTGAAATTTTGAATTTCTGAACATTTTGAAGATTTTGA 4920
4861 ATTAATTTTGAAGATTTGAAATTTTGAATTTCTGAACATTTTGAAGATTTTGA 4920
4921 CTTTAAAGGTAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATA 4980
4921 CTTTAAAGGTAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATA 4980
4981 AAAGAAATCCGAGAAAGCACTGGGAATGACATGGAATAACCCGCGCTCGCGCG 5040
4981 AAAGAAATCCGAGAAAGCACTGGGAATGACATGGAATAACCCGCGCTCGCGCG 5040
5041 ACTGTAAAGCTATAGTAGAGCGGCGCAAGCTCGTGTCTCATCATACCTGTTGGA 5100
5041 ACTGTAAAGCTATAGTAGAGCGGCGCAAGCTCGTGTCTCATCATACCTGTTGGA 5100
5101 AACCCGCAATTCGTTGCACTATGCGGGAATAGGCTTTTCAAGAGCTCTGCTTCC 5160
5101 AACCCGCAATTCGTTGCACTATGCGGGAATAGGCTTTTCAAGAGCTCTGCTTCC 5160
5161 GGTATAGGTGATTTGACACACCTCTCTCACTTGGGCGAGCTATTAATCTTTTTC 5220
5161 GGTATAGGTGATTTGACACACCTCTCTCACTTGGGCGAGCTATTAATCTTTTTC 5220
5221 TTCTTTGACCTCACTTACTAGGCAAGTTTATTTTGAAGGACCAACCGGTTTGT 5280
5221 TTCTTTGACCTCACTTACTAGGCAAGTTTATTTTGAAGGACCAACCGGTTTGT 5280

5281 GAAGTCTAGAAATCAACCAATTTTGGGAAGCTTCTAGAAGCTATGAATGTTCTTT 5340
5281 GAAGTCTAGAAATCAACCAATTTTGGGAAGCTTCTAGAAGCTATGAATGTTCTTT 5340
5341 TGAACATGATTAATTTGTTGTTTCTTTTCAAAATGCAACATCTTTTCAAAATCAT 5400
5341 TGAACATGATTAATTTGTTGTTTCTTTTCAAAATGCAACATCTTTTCAAAATCAT 5400
5401 GATTTTGTGAACCTGATGATTTTGAATCCGATTTTTCCTAAATCCGTTT 5460
5401 GATTTTGTGAACCTGATGATTTTGAATCCGATTTTTCCTAAATCCGTTT 5460
5461 GAAAAAATCTGACCTTTCCGAAATTAATGAACATTTGTTGCAAGATGATCTCT 5520
5461 GAAAAAATCTGACCTTTCCGAAATTAATGAACATTTGTTGCAAGATGATCTCT 5520
5521 TTTCAATGAGCGATTTTCTTAAATATCAACATTTTCTATTCATTAAGCTTTCC 5580
5521 TTTCAATGAGCGATTTTCTTAAATATCAACATTTTCTATTCATTAAGCTTTCC 5580
5581 TTTTAATGAGCACTATCTTAAGATTTGTTGACCTTTTAAATTTCTTAAATATGA 5640
5581 TTTTAATGAGCACTATCTTAAGATTTGTTGACCTTTTAAATTTCTTAAATATGA 5640
5641 TTTTCTTCAAAAGCAACGTTTAAACGTTGACCGCTGAACCAACCAACCAACCGGGA 5700
5641 TTTTCTTCAAAAGCAACGTTTAAACGTTGACCGCTGAACCAACCAACCAACCGGGA 5700
5701 AACCATGACCTGCTGAACAGGCGAGGCTTTCAATGATTTGGTGTCTAATACAGCG 5760
5701 AACCATGACCTGCTGAACAGGCGAGGCTTTCAATGATTTGGTGTCTAATACAGCG 5760
5761 CCCGTAATTAACGAAAGGAAATTTGAACCAACCACTACTAGTGAATTAATG 5820
5761 CCCGTAATTAACGAAAGGAAATTTGAACCAACCACTACTAGTGAATTAATG 5820
5821 AATATCAAGATTAATAAGGGAATAAATACTAATCCCTGAATTCCTGTTCTAAT 5880
5821 AATATCAAGATTAATAAGGGAATAAATACTAATCCCTGAATTCCTGTTCTAAT 5880
5881 ATTTGTTTGGGAGAACTAATCTGAAGAACTAATCTAATTTCCGCAATTAATAATA 5940
5881 ATTTGTTTGGGAGAACTAATCTGAAGAACTAATCTAATTTCCGCAATTAATAATA 5940
5941 TTATGATTCGGGGAGATTAATTAATTAATCAAGATCAACCAAGATGCTCTCAAGAAA 6000
5941 TTATGATTCGGGGAGATTAATTAATTAATCAAGATCAACCAAGATGCTCTCAAGAAA 6000
6001 ACCCAAGAAAGTCTAGATTTTGTCTCAAGACCGAAGATAGATGATGATTTCTGA 6060
6001 ACCCAAGAAAGTCTAGATTTTGTCTCAAGACCGAAGATAGATGATGATTTCTGA 6060
6061 ATTAGTTCACCTTTTCCCAAGATTTGAAGAAAGATGATGATTTCTGAATTTGGTGG 6120
6061 ATTAGTTCACCTTTTCCCAAGATTTGAAGAAAGATGATGATTTCTGAATTTGGTGG 6120
6121 GAGATATCATTTCTGAAATTAAGTCAATTTGTTTCCGCGCACAGCAACCCCGAGTGG 6180
6121 GAGATATCATTTCTGAAATTAAGTCAATTTGTTTCCGCGCACAGCAACCCCGAGTGG 6180
6181 AGCCGAAATTTGAATTAAGTGGTGGAGTCCAGAAAGATGATGATGATTTCTCAAGAG 6240
6181 AGCCGAAATTTGAATTAAGTGGTGGAGTCCAGAAAGATGATGATGATTTCTCAAGAG 6240
6241 TGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6300
6241 TGTAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6300
6301 GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6360
6301 GACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6360
6361 TTGATGTGCGAGCGAGTGGGAGCGGAGCCGCTATTAAGACCGGCCAAGCACCGC 6420

Db 6361 TTGATGTGGCAGGCGAGTGGGACGCGGACCGACTTAAAGCAGCGCCACGACCGC 6420
Qy 6421 CGCCGTTCTCAATCCATCCATCCCTTAGCTGATTTGATTTGACTAGTACTGATTCCTG 6480
Db 6421 CGCCGTTCTCAATCCATCCATCCCTTAGCTGATTTGATTTGACTAGTACTGATTCCTG 6480
Qy 6481 CCACATGCTAGTACTCCCTCTCTGTTTCTGTTGCGAATGTTACACAGAGCAACGCGCA 6540
Db 6481 CCACATGCTAGTACTCCCTCTCTGTTTCTGTTGCGAATGTTACACAGAGCAACGCGCA 6540
Qy 6541 CGGCGAGGCG 6600
Db 6541 CGGCGAGGCG 6600
Qy 6601 CGGCGAAGCAACG 6660
Db 6601 CGGCGAAGCAACG 6660
Qy 6661 GGAACGCGCATCTGCG 6720
Db 6661 GGAACGCGCATCTGCG 6720
Qy 6721 CAGCGCGAGCGTGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6780
Db 6721 CAGCGCGAGCGTGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6780
Qy 6781 GTCCGTTTCCCG 6840
Db 6781 GTCCGTTTCCCG 6840
Qy 6841 GGGCAGCG 6900
Db 6841 GGGCAGCG 6900
Qy 6901 ACATTTACAGCTTCAACCGTAATGTAAGCGTGAAGATGATGCGCGGTTTACTAGTGC 6960
Db 6901 ACATTTACAGCTTCAACCGTAATGTAAGCGTGAAGATGATGCGCGGTTTACTAGTGC 6960
Qy 6961 CGGCGCGCTGTTCTTCCCG 7020
Db 6961 CGGCGCGCTGTTCTTCCCG 7020
Qy 7021 ACAACAGCGCGCGTGAAGAGAGCGTGTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCG 7080
Db 7021 ACAACAGCGCGCGTGAAGAGAGCGTGTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCG 7080
Qy 7081 GACGCTTCTCTCAACCG 7140
Db 7081 GACGCTTCTCTCAACCG 7140
Qy 7141 CAGACTGCGCGCGCGCAACATAGTCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7200
Db 7141 CAGACTGCGCGCGCGCAACATAGTCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7200
Qy 7201 GGGGATTTCAAAAGCTGGAAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7260
Db 7201 GGGGATTTCAAAAGCTGGAAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7260
Qy 7261 ATCGACATCGACTCGCTGGAATTCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7320
Db 7261 ATCGACATCGACTCGCTGGAATTCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7320
Qy 7321 AACCCAAACATTCGCTGCG 7380
Db 7321 AACCCAAACATTCGCTGCG 7380
Qy 7381 CCATGATCTCTGCTGCTGCTGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7440
Db 7381 CCATGATCTCTGCTGCTGCTGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7440
Qy 7441 TTGCTAATCTGTGTGATGATGCTGTTGATTACAGGTGCGCGAGGTGCGAAGAAC 7500
Db 7441 TTGCTAATCTGTGTGATGATGCTGTTGATTACAGGTGCGCGAGGTGCGAAGAAC 7500

Db 7441 TTGCTAATCTGTGTGATGATGCTGTTGATTACAGGTGCGCGAGGTGCGAAGAAC 7500
Qy 7501 TCGGAATATGTTGATTCGCTGACGAGGTTTACGCGAACTGTTTCTGCGACGCGCGCG 7560
Db 7501 TCGGAATATGTTGATTCGCTGACGAGGTTTACGCGAACTGTTTCTGCGACGCGCGCGCG 7560
Qy 7561 TTATCCGATGAGGCGCTTTTGGGCGCAATTGCGCGCGCGCTGTGTCATTTGATCTGTCGA 7620
Db 7561 TTATCCGATGAGGCGCTTTTGGGCGCAATTGCGCGCGCGCTGTGTCATTTGATCTGTCGA 7620
Qy 7621 AGTCGTGATGATGCTGATGAGCGACTTGAATGGGTGCGCGGTGACGACCCCAAGAGA 7680
Db 7621 AGTCGTGATGATGCTGATGAGCGACTTGAATGGGTGCGCGGTGACGACCCCAAGAGA 7680
Qy 7681 TTTTGAAGAAAATGAAGTACTTGAATGCTTCCCTATCATTTCTCATATGCTACTGTGGG 7740
Db 7681 TTTTGAAGAAAATGAAGTACTTGAATGCTTCCCTATCATTTCTCATATGCTACTGTGGG 7740
Qy 7741 GATTAGTATTTTGTGTAATTTGTAATGCTGCTTGTGTTATTCAGATCTCTAGCTTATAC 7800
Db 7741 GATTAGTATTTTGTGTAATTTGTAATGCTGCTTGTGTTATTCAGATCTCTAGCTTATAC 7800
Qy 7801 GAATTAACCTTAATGCTGACAGGACCGAGCACTTCTGTCAGTTAGTCTTGTGTTCTT 7860
Db 7801 GAATTAACCTTAATGCTGACAGGACCGAGCACTTCTGTCAGTTAGTCTTGTGTTCTT 7860
Qy 7861 GCCCTATTTTGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7920
Db 7861 GCCCTATTTTGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7920
Qy 7921 AGAGTTTGTGTAATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7980
Db 7921 AGAGTTTGTGTAATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7980
Qy 7981 TTTAGTCAATCCCTTGTGTTGTTGATTCATTAATAATAATAACAGAACCGAGAAC 8040
Db 7981 TTTAGTCAATCCCTTGTGTTGTTGATTCATTAATAATAATAACAGAACCGAGAAC 8040
Qy 8041 AACTATGTAATGAATATTTTCTTGAAGAAATATACATTTGATGCGATGCGATAT 8100
Db 8041 AACTATGTAATGAATATTTTCTTGAAGAAATATACATTTGATGCGATGCGATAT 8100
Qy 8101 TCTTTGATGATGAATGAATATTTTGAAGAAATATTTGAAGAAATATTTGAAGAAAT 8160
Db 8101 TCTTTGATGATGAATGAATATTTTGAAGAAATATTTGAAGAAATATTTGAAGAAAT 8160
Qy 8161 GTTAAGAAATTTCTTACCTACAAACCTTGTCTGCACTTCTCATATTTTGAATATCTTGA 8220
Db 8161 GTTAAGAAATTTCTTACCTACAAACCTTGTCTGCACTTCTCATATTTTGAATATCTTGA 8220
Qy 8221 CATTAATGTTGCTTATGATTCACAAACCTTAATATGATTTTGTGCTATCAAAATGT 8280
Db 8221 CATTAATGTTGCTTATGATTCACAAACCTTAATATGATTTTGTGCTATCAAAATGT 8280
Qy 8281 TTAGGAAGCTGCTCTTAAATTTCTTGAAGAAACAAAGAGATTTCTTAAAGAGATAT 8340
Db 8281 TTAGGAAGCTGCTCTTAAATTTCTTGAAGAAACAAAGAGATTTCTTAAAGAGATAT 8340
Qy 8341 TGTCTACTAAAGAAATCATGAGATATGTTATAGGAAATTAAGAAACAAATATAT 8400
Db 8341 TGTCTACTAAAGAAATCATGAGATATGTTATAGGAAATTAAGAAACAAATATAT 8400
Qy 8401 TACGTGCTCTCAAGCGCAAGAGATGATGTTTGTATATGTAAGTGAAGCTTACACTTA 8460
Db 8401 TACGTGCTCTCAAGCGCAAGAGATGATGTTTGTATATGTAAGTGAAGCTTACACTTA 8460
Qy 8461 CTTTAAAGTAAATCTGGAATCTGAGTCAATCCAAACAAATCAATCAAAATATATAT 8520
Db 8461 CTTTAAAGTAAATCTGGAATCTGAGTCAATCCAAACAAATCAATCAAAATATATAT 8520
Qy 8521 TATGTTTGTGTAATGATCTTTTGAAGATGATGATTTGAAGATTAATGAAGAGTTG 8580
Db 8521 TATGTTTGTGTAATGATCTTTTGAAGATGATGATTTGAAGATTAATGAAGAGTTG 8580

8581 AAATATTTAGGACCTAATCTTCTGATATCATTTGTCATTTTGTATTAATCTGAA 8640
8581 AATATTTTGAAGACTAATCTTCTGATATCATTTGTCATTTTGTATTAATCTGAA 8640
8641 ATGGTAACACTCAATCATATTTACAAAAGTTTCTCCCATTTTGTATTAATGACTT 8700
8641 ATGGTAACACTCAATCATATTTACAAAAGTTTCTCCCATTTTGTATTAATGACTT 8700
8701 CTTTCTAATACCAATGATTAATCTTCATGTAAGAGGTCAACTTAATCAATCTTT 8760
8701 CTTTCTAATACCAATGATTAATCTTCATGTAAGAGGTCAACTTAATCAATCTTT 8760
8761 GGAGAGATCCATGACGACATTAATTTTGTGCAAGCTCGCAAGAGAAAGATCTGAT 8820
8761 GGAGAGATCCATGACGACATTAATTTTGTGCAAGCTCGCAAGAGAAAGATCTGAT 8820
8821 TTTATGTCAGGTAGGAATGATATGATGAGCAATTTTAAAGAAAATGATGAAATATAT 8880
8821 TTTATGTCAGGTAGGAATGATATGATGAGCAATTTTAAAGAAAATGATGAAATATAT 8880
8881 ATCTCTGTTATATTAACAATATCTTCTCATCTCAATAAATGATCTTAATCACTTAC 8940
8881 ATCTCTGTTATATTAACAATATCTTCTCATCTCAATAAATGATCTTAATCACTTAC 8940
8941 ACAATTTTATACTAATCTAGTACAAAGTTGAAACAGTTATTTTGGGACAGAGGAGTAG 9000
8941 ACAATTTTATACTAATCTAGTACAAAGTTGAAACAGTTATTTTGGGACAGAGGAGTAG 9000
9001 TATATATGTTGTGAGAACATTAAGTTATGTTGACTGATATATGCTTCTTAATATGAAA 9060
9001 TATATATGTTGTGAGAACATTAAGTTATGTTGACTGATATATGCTTCTTAATATGAAA 9060
9061 CATGTCCTTAATGTTTGTGATGATACGAAGTTCTTAATAGTTTCGAGATGACAC 9120
9061 CATGTCCTTAATGTTTGTGATGATACGAAGTTCTTAATAGTTTCGAGATGACAC 9120
9121 ACATTAATGATTAACATATCATATGTCAGAAAATGATTAACCAATAGATATCTTCT 9180
9121 ACATTAATGATTAACATATCATATGTCAGAAAATGATTAACCAATAGATATCTTCT 9180
9181 TTTTATGCAAGACATGACGATGATCTTCTGTAATCTTATGCTTTTCTTTTTC 9240
9181 TTTTATGCAAGACATGACGATGATCTTCTGTAATCTTATGCTTTTCTTTTTC 9240
9181 TTTTATGCAAGACATGACGATGATCTTCTGTAATCTTATGCTTTTCTTTTTC 9240
9241 TCGTACATGTTTGTGCTTCTGCAAAAATATATTAACCAAGCAATGTCGAATGAT 9300
9241 TCGTACATGTTTGTGCTTCTGCAAAAATATATTAACCAAGCAATGTCGAATGAT 9300
9301 ATTAATTAATTTGAGGTGTTTTCACCAACTTAATATCTTCAATAGTTCTTAATAAAC 9360
9301 ATTAATTAATTTGAGGTGTTTTCACCAACTTAATATCTTCAATAGTTCTTAATAAAC 9360
9361 GATATATGTTTGAAGTCTTAACCAAACTTAATATGTTTCTCTTAATACAGGAGTGT 9420
9361 GATATATGTTTGAAGTCTTAACCAAACTTAATATGTTTCTCTTAATACAGGAGTGT 9420
9421 TCTTGAATGAAAAATGAGTCCGATATTAATCTTGTGCTGCTGCTTCTTCAAGA 9480
9421 TCTTGAATGAAAAATGAGTCCGATATTAATCTTGTGCTGCTGCTTCTTCAAGA 9480
9481 TGGACTCGAAGAGGTCAATCTTCTGCAAGAAAGAAAGAAAGAAATCTTAATATG 9540
9481 TGGACTCGAAGAGGTCAATCTTCTGCAAGAAAGAAAGAAAGAAATCTTAATATG 9540
9481 TGGACTCGAAGAGGTCAATCTTCTGCAAGAAAGAAAGAAAGAAATCTTAATATG 9540
9541 TTTGTAATGTTTACACACCCCTAGTGTATCTGACTGAAAGTGTAAATCTTCTTACT 9600
9541 TTTGTAATGTTTACACACCCCTAGTGTATCTGACTGAAAGTGTAAATCTTCTTACT 9600
9601 ATCCCATTTATATTTCAATTAACATATGATGATGTTGTTGTTGCTGCTCAAGT 9660
9601 ATCCCATTTATATTTCAATTAACATATGATGATGTTGTTGTTGCTGCTCAAGT 9660

9661 CATGTAATCTACTTTTGAATGATATGAGCCATATGCTTGCATCAAGTTCAATTAATAAT 9720
9661 CATGTAATCTACTTTTGAATGATATGAGCCATATGCTTGCATCAAGTTCAATTAATAAT 9720
9721 GGTGTGTACCAATGATGATGATGAGGCGAGGTGTTTGAACCACTTTTCAACAAAAT 9780
9721 GGTGTGTACCAATGATGATGATGAGGCGAGGTGTTTGAACCACTTTTCAACAAAAT 9780
9781 CTATATCTTTCAACAAATGAAACCTTGAGTCCCTTGAGTGAAGTCAATCACTCTT 9840
9781 CTATATCTTTCAACAAATGAAACCTTGAGTCCCTTGAGTGAAGTCAATCACTCTT 9840
9841 GAATATGCTATGATTTCCATGCTGATGATGAAATGATGAAATGATGAAATGATG 9900
9841 GAATATGCTATGATTTCCATGCTGATGATGAAATGATGAAATGATGAAATGATG 9900
9901 CATGCAAGTTTATTAATGTTTATTTCAATTAATGAGAACCTTGAATATCTTCTAGAC 9960
9901 CATGCAAGTTTATTAATGTTTATTTCAATTAATGAGAACCTTGAATATCTTCTAGAC 9960
9961 ACATTTCTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10020
9961 ACATTTCTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10020
10021 GGGGGGGGGTGAATAGCGGTTTATTAACCAATGATATTTGAGAAATCTTAATGTTGA 10080
10021 GGGGGGGGGTGAATAGCGGTTTATTAACCAATGATATTTGAGAAATCTTAATGTTGA 10080
10081 ATTAATCTAGTGAATATTTTTCATTAATTAAGGTTGCTTTATGATCTCAATGTTACA 10140
10081 ATTAATCTAGTGAATATTTTTCATTAATTAAGGTTGCTTTATGATCTCAATGTTACA 10140
10141 TCAAGGATATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10200
10141 TCAAGGATATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10200
10201 CCAAC 10260
10201 CCAAC 10260
10261 GTCATACAGATCAAACTAATGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 10320
10261 GTCATACAGATCAAACTAATGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 10320
10321 AAAACCGTACGACATATGATGATGATGATGATGATGATGATGATGATGATGATG 10380
10321 AAAACCGTACGACATATGATGATGATGATGATGATGATGATGATGATGATGATG 10380
10381 CTGCGAGAAAAGTTCTATTAATTAAGTGGGCTTGCAGAAAGAAAGAGCGTCAAGATTGC 10440
10381 CTGCGAGAAAAGTTCTATTAATTAAGTGGGCTTGCAGAAAGAAAGAGCGTCAAGATTGC 10440
10381 CTGCGAGAAAAGTTCTATTAATTAAGTGGGCTTGCAGAAAGAAAGAGCGTCAAGATTGC 10440
10441 CATATGAGATCAACCACTTAAGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAG 10500
10441 CATATGAGATCAACCACTTAAGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAG 10500
10501 AGCAAACTCCGAGTATGCTTTATTAAGGCTTATGAGGCTTATGAGGCTTATGAGGCTT 10560
10501 AGCAAACTCCGAGTATGCTTTATTAAGGCTTATGAGGCTTATGAGGCTTATGAGGCTT 10560
10561 TTAATGCAATTAATGACGACCTTATGATGATGATGATGATGATGATGATGATGATG 10620
10561 TTAATGCAATTAATGACGACCTTATGATGATGATGATGATGATGATGATGATGATG 10620
10621 AGTACATCAATGAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10680
10621 AGTACATCAATGAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 10680
10681 GTCCATATGCGAGTAAATGACATGAGCGGCTGATGATGATGATGATGATGATGATGAT 10740
10681 GTCCATATGCGAGTAAATGACATGAGCGGCTGATGATGATGATGATGATGATGATGAT 10740
10741 GCTGTGCGCCACCACTGACGAGATGTTTATGAGGAGGAGGAGGAGGATGAGCTCCACCA 10800
10741 GCTGTGCGCCACCACTGACGAGATGTTTATGAGGAGGAGGAGGAGGATGAGCTCCACCA 10800

[illegible][illegible]

OY 1717 GCACATGCACTGCGTGAATCATGCGGACAGAAACCAACCCGCTGCTCATCATATAA 1776
 DB 960 CGACATGCACTGCGTGAATCATGCGGACAGAAACCAACCCGCTGCTCATCATATAA 1019
 OY 1777 CCCCACAAACCCGTCGCGGACGCTTACTCTTACGACCATCTGTCCAAAGTTTCACATCC 1836
 DB 1020 CCCCACAAACCCGTCGCGGACGCTTACTCTTACGACCATCTGTCCA----- 1066
 OY 1837 TTTCCTTCTGCAATATGATTCAGTTCAGTGCACCTGCTGAATCTTTTCCCATGCC 1896
 DB 1067 ----- 1066
 OY 1897 ATACTGACTGATGTTGCTCAATTAGTGCAGGAGTGGGAAAAGGCTCGAATATTGAT 1956
 DB 1067 -----AGTGCAGAGAGTGGGAAAAGGCTCGAATATTGAT 1103
 OY 1957 GATTGCTGACGAGATATACGCGCAAGCTGTCTGGGACAGCCCGCTTATCCCATGAG 2016
 DB 1104 GATTGCTGACGAGATATACGCGCAAGCTGTCTGGGACAGCCCGCTTATCCCATGAG 1163
 OY 2017 AGTGTTCGGGACATACACCCCTGTCTGTCTCCATAGGTTCTCTTCAAGTCATGATAGT 2076
 DB 1164 AGTGTTCGGGACATACACCCCTGTCTGTCTCCATAGGTTCTCTTCAAGTCATGATAGT 1223
 OY 2077 GCCTGATGCGGCTTGGATGAGTGGTACGAGGCTGACACCCAGAAAGATCTTACAGAAAG 2136
 DB 1224 GCCTGATGCGGCTTGGATGAGTGGTACGAGGCTGACACCCAGAAAGATCTTACAGAAAG 1283
 OY 2137 TAAGGTACTTAATC 2151
 DB 1284 TAAGATCTCTACATC 1298

RESULT 3

AAV48147 standard; cDNA to mRNA; 1660 BP.

AAV48147;

 DT 27-AUG-2003 (revised)
 DT 27-OCT-1998 (first entry)

DE Nicotianamine aminotransferase 49564.15 molecular weight protein, gene.

 KM ds; nicotianamine aminotransferase; plant; iron absorption;
 KM iron deficiency chlorosis.

XX Poaceae.

XX Location/Qualifiers

XX Key 62.1447

XX /tag= a

XX /product= "Nicotianamine aminotransferase"

XX EP860499-A2.

XX 26-AUG-1998.

XX 19-FEB-1998; 98EP-00102891.

XX 21-FEB-1997; 97JP-00037499.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX Mori S, Nakanishi H, Takahashi M,

XX MPI; 1998-439341/38.

XX P-PsDB; AAW61642.

 PT New nicotianamine aminotransferase protein and DNA - useful for enhancing
 PT iron absorption of plant cells.

PS Claim 4; Page 12-13; 17pp; English.

 CC The nicotianamine aminotransferase can be used in a plasmid to transform
 CC plant cells to produce cells with enhanced iron absorption, and it is
 CC implied (though not stated) that plants with improved resistance to iron
 CC deficiency chlorosis in calcareous soils can be regenerated from the
 CC transformed cells. The gene fragment can be used to detect, amplify
 CC and/or isolate nicotianamine aminotransferase genes. (Updated on 27-AUG-
 CC 2003 to correct OS field.)

XX Sequence 1660 BP; 423 A; 442 C; 430 G; 365 T; 0 U; 0 Other;

SQ Query Match 6.8%; Score 744.6; DB 2; Length 1660;

Best Local Similarity 80.6%; Pred. No. 1e-115;

Matches 1005; Conservative 0; Mismatches 4; Indels 238; Gaps 2;

OY 6457 ATTGACTGAGTGAATGTTCAATTCCTGCGCACTGCTGATGATCTCTCTCTGTTCTGTCGCG 6516
 DB 1 ATTGACTGAGTGAATGTTCAATTCCTGCGCACTGCTGATGATCTCTCTCTGTTCTGTCGCG 60
 OY 6517 AATGTTACACGAGCAACGAGCGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6576
 DB 61 AATGTTACACGAGCAACGAGCGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 OY 6577 CAACGCGCACG 6636
 DB 121 CAACGCGCACG 180
 OY 6637 GGAATGGAATTTCCCG 6686
 DB 181 GGAATGGAATTTCCCG 240
 OY 6697 CATCGCGCGCGATGAGTAAAGATCAAGATCAGCGAGCGAGTGGAGAGAGCGCGCGCGCGCGCG 6756
 DB 241 CATCGCGCGCGATGAGTAAAGATCAAGATCAGCGAGCGAGTGGAGAGAGCGCGCGCGCGCG 300
 OY 6757 GCTGCGCGTGGCCACGAGTGAACCGGTCGTTCCCGGCTTCCGACAGCGCGCGTGAAGC 6816
 DB 301 GCTGCGCGTGGCCACGAGTGAACCGGTCGTTCCCGGCTTCCGACAGCGCGCGTGAAGC 360
 OY 6817 CGAGAGAGCGCGTGGCG 6876
 DB 361 CGAGAGAGCGCGTGGCG 420
 OY 6877 CGGCTCCCGCGCGCGAGAGGTAACTTACAGCTTACCGGTATGATGCGTGAAGAT 6936
 DB 421 CGGCTCCCGCGCGCGAGAGGTAACTTACAGCTTACCGGTATGATGCGTGAAGAT 439
 OY 6937 GCATGCGCGGTTTACTTAAGTCGTCGCGCGCGCTGTTCTTCCCGGTGCGTTCAAAATTTTA 6996
 DB 440 ----- 439
 OY 6997 ACCTTTTAATAGTACTTTATTAATAAACAACAGCGCGCGTGAAGAGCACTTGTCAAGGCG 7056
 DB 440 -----AGCGCGTGAAGAGCACTTGTCAAGGCG 469
 OY 7057 GTGCGCTTAAGCTATCGGCGAGAGAGTCTTCTCAACGCGCGCGAGAACTCAGGCGATC 7116
 DB 470 GTGCGCTTAAGCTATCGGCGAGAGAGTCTTCTTCTCAACGCGCGCGAGAACTCAGGCGATC 529
 OY 7117 GAAGTCAATATCCGAGTGTGCGCGAGAGTGCAGCGCGCGCAATATGCTTCCCGGCGCA 7176
 DB 530 GAAGTCAATATCCGAGTGTGCGCGAGAGTGCAGCGCGCGCAATATGCTTCCCGGCGCA 589
 OY 7177 GGCATTCGAATTTACGAGCG 7236
 DB 590 GGCATTCGAATTTACGAGCG 649
 OY 7237 CTGATCCCGCGCAAGGAGTGGAGATGAGCATGAGTCTGATGATCATGCGCGCAAG 7296
 DB 650 CTGATCCCGCGCAAGGAGTGGAGATGAGCATGAGTCTGATGATCATGCGCGCAAG 709
 OY 7297 AACACACCGGAGTGTATCATTAACCAACCAATCGTGGCGAGCGTTTACTCTTAC 7356


```

Db 710 AACACACCGCATGATCATCAATAACCAACAAATCCGTGCGGAGGCTTACTCTAC 769
Qy 7357 GACCATTCGGCCCAAGTTTGCATCCATCCTCTGCTGTTGATGACCGCTCTCT 7416
Db 770 GACCATTCGGCCA----- 782
Qy 7417 TTGAACATAGTATGATGATGCTTGGTTGTAATCGTGTGATGATGCTTTGTTATC 7476
Db 783 ----- 782
Qy 7477 AGGTGCGGAGGTGCGCAAGAAAGCTCGAATAATGCTGATCGCTGACGAGTTTACGGCA 7536
Db 783 AGGTGCGGAGGTGCGCAAGAAAGCTCGAATAATGCTGATCGCTGACGAGTTTACGGCA 842
Qy 7557 AACTGCTTTCGGGAGCGCCCGCTTATCCCGATGGGCTCTTTGGGCAATTCGCCCGG 7596
Db 843 AACTGCTTTCGGGAGCGCCCGCTTATCCCGATGGGCTCTTTGGGCAATTCGCCCGG 902
Qy 7597 TCTTGTCCATTGATCTCTGTCCAAGTGTGATGATGCTGATGCGACTTGGATGGG 7656
Db 903 TCTTGTCCATTGATCTCTGTCCAAGTGTGATGATGCTGATGCGACTTGGATGGG 962
Qy 7657 TGGCGGTGACGACCCCAAAAGATTTTGAAGAAACTAAGTAGCT 7703
Db 963 TGGCGGTGACGACCCCAAAAGATTTTGAAGAAACTAAGTAGCT 1009

```

RESULT 4
ADK52801
ID ADK52801 standard; DNA; 568 BP.

ADK52801;

06-MAY-2004 (first entry)

Plant DNA sequence which confers altered metabolic characteristic #184.

altered metabolic characteristic; plant; acid metabolism;
alcohol metabolism; fatty acid metabolism;
branched fatty acid metabolism; alkaloid metabolism;
amino acid metabolism; ester metabolism; glyceride metabolism;
phenolic metabolism; carbohydrate metabolism; steroid metabolism;
terpene metabolism; isoprenoid metabolism; alkene metabolism;
alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.

Oryza sativa.

WO2003020936-A1.

13-MAR-2003.

30-AUG-2002; 2002WO-US027884.

31-AUG-2001; 2001US-0316471P.

(DOMC) DOM CHEM CO.

(DOMC) DOM AGRSCIENCES LLC.

Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ,

Oriado JVB, Crooley R, Reddy AS, Shukla V, Larrina I, Miller BA,

WPI; 2003-313091/30.

Novel genes that confer altered metabolic characteristics in Nicotiana
benhamiana plants, useful for altering the levels of metabolites e.g.
acids, fatty acids, amino acids, carbohydrates, hydrocarbons and steroids.

Claim 1; SEQ ID NO 184; 2576bp; English.
The invention comprises DNA sequences which confer an altered metabolic
characteristic when they are expressed in a plant. The DNA sequences of

CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC altered glyceride metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered steroid, oxygenated terpene, or
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.

Sequence 568 BP; 113 A; 201 C; 171 G; 83 T; 0 U; 0 Other;

Query Match 1.7%; Score 184.4; DB 10; Length 568;
Best Local Similarity 78.9%; Pred. No. 26-21;
Matches 232; Conservative 0; Mismatches 60; Indels 2; Gaps 1;

```

Qy 1008 CGCGCGCGCGCAGAGAGAGAGAGC--GGTGAAGTGAATTTCGCGGTGCCAAGA 1065
Db 64 CGCGCGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
Qy 1066 CGCGGTCTGCGCGCGAGCGGGGCGCAACATGACATCCGGGCGATACGATCAAGATCAG 1125
Db 124 CGCGCGAGATGCGCGCGCGCGCGCGGACAGATGAGATCCGGGCGGTGCGGTACAAAGATCAG 183
Qy 1126 CGCGAGCGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
Db 184 CGCGAGCGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
Qy 1186 CGTGTTCCTCCGCTTCCGCAAGCGCGCGTGAAGCGAGAGAGCGTGCAGCGCGCTGCG 1245
Db 244 CGTGTTCCTCCGCTTCCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCG 303
Qy 1246 CACCGCGCAATTCAATCTGCTACCCCGCGCGCGTGCCTCCCGCGCGCAGAG 1299
Db 304 CACCGCGCAATTCAATCTGCTACCCCGCGCGCGTGCCTCCCGCGCGCAGAG 357

```

RESULT 5
ADA70556

ID ADA70556 standard; DNA; 1233 BP.

ADA70556;

20-NOV-2003 (first entry)

Rice gene, SEQ ID 3879.

Plant; bacterial infection; fungal infection; viral infection; rice;

gene; ds.

Oryza sativa.

WO200300898-A1.

03-JAN-2003.

22-JUN-2001; 2001WO-IB001105.

22-JUN-2001; 2001WO-IB001105.

(SYGN) SYNGENTA PARTICIPATIONS AG.

Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,

Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G,

WPI; 2003-175290/17.

Identifying at least one gene involved in plant resistance or response to
pathogenic infection for conferring resistance or tolerance to a plant to
bacterial, fungal or viral infection by determining or detecting plant
gene expression.

XX Claim 6; SEQ ID NO 3879; 899bp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC the expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

SO Sequence 1233 BP; 221 A; 397 C; 412 G; 201 T; 0 U; 2 Other;

Query Match 1.3%; Score 147; DB 8; Length 1233;

Best Local Similarity 67.0%; Pred. No. 4.9e-15;

Matches 223; Conservative 1; Mismatches 106; Indels 3; Gaps 1;

QY 1480 CAGGCGGTGGCAGACCTGTGCGAGGGGTGCGGACATGCTATGCGCGACGACGT 1539

DB 294 CAGGCGGTGGCAGACCTGTGCGAGGGGTGCGGACCTCTTCCGCGGTGCGCGGACGT 353

QY 1540 CTTCCTCACCGCCGCGGACCCAGCGCATCGAGGTCAATATCCCGGTGCTGCGCCAGAC 1599

DB 354 GGTGCTACCGCCGCGGTCAACAGCGCGTCAAGATCATGATGTCGCTCGC---GTC 410

QY 1600 CGCGGCGCCCAACATTGCTGCTCCCAAGCCAGGCTAACCAACTACGAGCGCGCGCGC 1659

DB 411 GCGCGGCGCCCAACGCTGCTGCTCCCGCGCGCGCTACCGCGTACGCGTGGCGCGCGC 470

QY 1660 GTTCAACAGCGGTGAGGCTGCGGCAATTTGACCTCATCCCGCAAGGGGTGGAGATGA 1719

DB 471 CTTGAGCGGCTTCGATTCGCGCACTTGTGACTCTCTCCCAACGCGGTGGAGGTGA 530

QY 1720 CATGCACTCGCTGATCCATCGCCGACAGAACACACACCGCATGTATCATTAACCC 1779

DB 531 CTTGCGCGGCGTCAAGCGCTCTGCGGACGCCAAMACGTCGCAATGATCGAGAGCGC 590

QY 1780 CAACAAACCGTGGCGGACGTTTACTCTTACGA 1812

DB 591 GAGGAGCTGGGATTAATGATGATGACGACGA 623

RESULT 6

AB214064 ID AB214064 standard; DNA; 1338 BP.

AC AB214064;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1869.

KW Arabidopsis thaliana; plant; gene; stress; transgenic; db.

OS Arabidopsis thaliana.

XX Arabidopsis thaliana.

PN MO200216655-A2.

PD 28-FEB-2002.

PF 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SGRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Krepe J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and
 CC producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 1869; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office

SO Sequence 1338 BP; 375 A; 307 C; 326 G; 330 T; 0 U; 0 Other;

Query Match 1.3%; Score 142.2; DB 6; Length 1338;

Best Local Similarity 63.1%; Pred. No. 3.2e-14;

Matches 219; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1481 AGCGCCCTGGCAGACCTGTGCGAGGGGTGCGGACATGCTATGCGCGACGACGTC 1540

DB 292 AGCGCGGTGGCTGAATTTTAAACGAGAACCTTCCGACGAAAGTGAAGCCGAGAGATGTG 351

QY 1541 TTCTCTACCGCCGCGGACCCAGCGCATCGAGGTCAATATCCCGGTGCTGCGCCAGAC 1600

DB 352 TATATCACCGGAGATTAACCAAGCATAGAGTCTGATGATTTCTTTCGCGGAAT 411

QY 1601 GCGGCGCCCAACATTGCTGCTCCCAAGCCAGGCTAACCAACTACGAGCGCGCGCGC 1660

DB 412 CCATCCGCCCAACATTACTTCCCAAGCCGCGGATATCTCTCACTGATGCTGCTGTC 471

QY 1661 TTCAACAGCGTGAAGGCTGCGGCAATTTGACCTCATCCCGCAAGGGGTGGAGATGAC 1720

DB 472 TATAGCGGCTCGAGATTGCAATTAATGATCTTCTCCCGAGATGATGGAAATCAAT 531

QY 1721 ATGCACTCGCTGATCCATCGCCGACAGAACACACACCGCATGTATCATTAACCC 1780

DB 532 CTCGATGGCTCGAGGCGCGCTGCGGATGAGATACCGTGGAAATGATATCAACCC 591

QY 1781 AACCAACCGTGGCGGACGTTTACTCTTACGACCATCTGTCAAGGT 1827

DB 592 AACCAATCATGTGGAACGTTCACTACCTACGACCATCTCAACAGGT 638

RESULT 7

ADG87698 ID ADG87698 standard; cDNA; 1338 BP.

AC ADG87698;

DT 22-APR-2004 (first entry)

DE A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #140.

KW Pathogen infection-related gene; plant; Peronospora parasitica;

KW defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;

KW oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.

OS Arabidopsis thaliana.

XX Arabidopsis thaliana.

PN MO200222675-A2.

PD 21-MAR-2002.

PF 14-SEP-2001; 2001WO-US028506.

XX 15-SEP-2000; 2000US-0232778P.
 PR 22-JUN-2001; 2001US-0300183P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA (UNIC-) UNIV NORTH CAROLINA.
 PA (GLAZ/) GLAZEBROOK J.
 PA (WANG/) WANG X.
 PA (DANG/) DANG J L.
 PA (EUG/) EUGEN T.
 PA (ZHU/) ZHU T.
 XX
 PI Glazebrook J, Wang X, Dang J, Eulgem T, Zhu T;
 DR WPI; 2002-292409/33.
 XX
 PT Novel isolated polynucleotide, useful for conveying pathogen resistance
 PS to plants, and for identifying plants infected with a pathogen.
 XX
 PS Claim 3: SEQ ID NO 140; 605bp; English.
 XX
 CC The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
 CC ADG87557) whose expression is altered in response to pathogen infection,
 CC and to homologues of these genes from other plants or fungi, especially
 CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
 CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
 CC expression of genes of the invention was upregulated or downregulated in
 CC Arabidopsis plants infected with the mycelia Peronospora parasitica,
 CC indicating that they play a role in defence mechanisms. The genes of the
 CC invention are regulated by RPP7 or RPP8 which act via unconventional
 CC signalling cascades, or by the RPP4-dependent pathway. The invention also
 CC relates to polypeptides encoded by the pathogen infection-related genes
 CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
 CC / expression cassettes, host cells and pathogen-resistant transgenic
 CC plants and their progeny comprising a polynucleotide of the invention;
 CC and a method of identifying a plant cell infected with a pathogen. The
 CC polynucleotide sequences and methods of the invention are useful for
 CC identifying plants infected with a pathogen, and for conferring
 CC resistance to pathogens such as mycelia, fungi, bacteria, viruses,
 CC nematodes and insects (e.g., aphids). The present sequence represents an
 CC Arabidopsis thaliana gene whose expression is altered in response to
 CC Peronospora parasitica infection. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WPIO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 1338 BP; 375 A; 307 C; 326 G; 330 T; 0 U; 0 Other;
 Query Match 1.3%; Score 142.2; DB 6; Length 1338;
 Best Local Similarity 63.1%; Pred. No. 3.2e-14;
 Matches 219; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
 Oy 1481 AGCGCGTGGAGACACCTGTGCGAGGCGGTCCGTCATCTATCGCCGACGACGTC 1540
 Db 292 AGGCGGTGGCTGATATTTAAACGAGAACTTCGACGAACTGAGGCGGAGATGTG 351
 Oy 1541 TTCCTACCGCGCGGAGACCCAGCGCATTCAGTCATATCCCGTGTGCGCCAGAC 1600
 Db 352 TATATCACCGGAGATGTAAACCAAGCCATAGAGATCGATGATTTCTTTCGCGAAAT 411
 Oy 1601 GCCCGCGCAACATTCTGCTCCCGAGCGCAGCTACCCAACTAGAGCGCGCGCGG 1660
 Db 412 CCATCCGCAACATTCTGCTCCCGAGCGCAGCTACCCAACTAGAGCGCGCGCGG 471
 Oy 1661 TTCAAAGGCTGAGAGTCCGCGCATTTGACCTCATCTCCCGACAGGGGTGGAGATGAC 1720
 Db 472 TATACCGCGCTCGAATTCGCAAAATACGATCTTCCCGAGAGATGGGAATCAAT 531
 Oy 1721 ATCGCTGCTGAGATTCATGCGCGCAAGAACACACCGCGCATGTATCAATAACCC 1780
 Db 532 CTCGATGCTTCGAGCGCGCTGCGAGATGAAATACCGTGCATATCAATCAACCC 591
 Oy 1781 AACACCGCTGCGGAGCGCTTACTCTACGACCATCTGTCCAAGT 1827

DB 592 AACATCCATGTGGAACGCTTACCTACCTACGACCATCTCAAGAGT 638
 RESULT 8
 ADA68509
 ID ADA68509 standard; DNA; 1338 BP.
 XX
 AC ADA68509;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Arabidopsis thaliana gene, SEQ ID 584.
 XX
 KM Plant; bacterial infection; fungal infection; viral infection; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN MO200300898-A1.
 XX
 PD 03-JAN-2003.
 XX
 PF 22-JUN-2001; 2001WO-1B001105.
 XX
 PR 22-JUN-2001; 2001WO-1B001105.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Qian S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
 DR WPI; 2003-175290/17.
 XX
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX
 PS Claim 6; SEQ ID NO 584; 899bp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 CC
 XX
 SQ Sequence 1338 BP; 375 A; 307 C; 326 G; 330 T; 0 U; 0 Other;
 Query Match 1.3%; Score 142.2; DB 8; Length 1338;
 Best Local Similarity 63.1%; Pred. No. 3.2e-14;
 Matches 219; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
 Oy 1481 AGCGCGTGGAGACACCTGTGCGAGGCGGTCCGTCATCTATCGCCGACGACGTC 1540
 Db 292 AGGCGGTGGCTGATATTTAAACGAGAACTTCGACGAACTGAGGCGGAGATGTG 351
 Oy 1541 TTCCTACCGCGCGGAGACCCAGCGCATTCAGTCATATCCCGTGTGCGCCAGAC 1600
 Db 352 TATATCACCGGAGATGTAAACCAAGCCATAGAGATCGATGATTTCTTTCGCGAAAT 411
 Oy 1601 GCCCGCGCAACATTCTGCTCCCGAGCGCAGCTACCCAACTAGAGCGCGCGCGG 1660
 Db 412 CCATCCGCAACATTCTGCTCCCGAGCGCAGCTACCCAACTAGAGCGCGCGCGG 471
 Oy 1661 TTCAAAGGCTGAGAGTCCGCGCATTTGACCTCATCTCCCGACAGGGGTGGAGATGAC 1720
 Db 472 TATACCGCGCTCGAATTCGCAAAATACGATCTTCCCGAGAGATGGGAATCAAT 531

OY 1721 ATCGACTGCTGATCCATCGCCGACAAACACACCAGCATGTCATCATTAACCCC 1780
DB 532 CTCGATGCGCTCGAGCGCGGCTCGATGAGATACCGTCGAACTGTCATCAACCCC 591
OY 1781 AACAAACCCGTCGCGAGCGTTTACTCTTACGACCACTCTGTCCAGGT 1827
DB 592 AACATCATGTGGAACGCTACCTACGACCATCTCAACAGGT 638

RESULT 9
AAC45930
ID AAC45930 standard; DNA; 1533 BP.
XX
AC AAC45930;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48285.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145152P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.

```
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148355P.
PR 13-AUG-1999; 99US-0148384P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151103P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157765P.
PR 06-OCT-1999; 99US-0157855P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160808P.
PR 22-OCT-1999; 99US-0160811P.
PR 22-OCT-1999; 99US-0160899P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 1.3%; Score 142.2; DB 3; Length 1533;
Best Local Similarity 63.1%; Pred. No. 3.3e-14;
Matches 219; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```
QY 1481 AGCGCGCTGCGAGAGCACTGTCGCGGCGGCTTACATCTATCGCCGACGACGTC 1540
DB 356 AGGCGCGTGGCTGAATATTAAACGAGAACTTCCGACCAAGCCGACGATGTG 415
```

```
QY 1541 TTCTCAACCGCGCGGACCCAGGCGATCGAGTCAATATCCCGGTCTGGCCGAC 1600
DB 416 TATATACCGGAGATGTATCAACGACATAGATGTATGATTTCTTGGCGGAAT 475
QY 1601 GCCGCGCAACATTGTGCTCCCGACGAGCTTACCAATACAGAGCGCGCGCG 1660
DB 476 CCATCCGCAACTTACTTCTCAAGCCGGGATATCTCACTACGATGCTGCTGTC 535
QY 1661 TTCAACAGGCTGAGGCTCGGCACTTTCGACCTTCATCCCGACAGGCGGATGAC 1720
DB 536 TATAGCGGCTCGAGATTCGCAATTCGATCTTCTCCGAGAGTATGGGAATCAAT 595
QY 1721 ATGCACTCGTGGATTCATTCGCGGCAAGAACACACGCGCATGGTCATTAACCCC 1780
DB 596 CTGATGCGCTTCGAGCGGCTGCGTGAATACCTGCGAATGGTAATCATCAACCC 655
QY 1781 AACCAACCGTGGCGGACGCTTACTCTCAACGACATCTGTCGACGCT 1827
DB 656 AACCAATCATGTGGAAGCTTACCTACGACGACATCTCAACGAGCT 702

RESULT 10
AAC35232
ID AAC35232 standard; DNA; 1535 BP.
XX
AC AAC35232;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9451.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0133484P.
PR 05-MAY-1999; 99US-0133485P.
PR 06-MAY-1999; 99US-0133486P.
PR 06-MAY-1999; 99US-0133487P.
PR 07-MAY-1999; 99US-0133486P.
PR 11-MAY-1999; 99US-0133425P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
```

PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144684P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 28-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148664P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156586P.
 PR 04-OCT-1999; 99US-0157717P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158028P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159548P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 1.3%; Score 142.2; DB 3; Length 1535;
 Best Local Similarity 63.1%; Pred. No. 3.3e-14;
 Matches 219; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1461 AGCCGCGGAGAGACCTGTGCGAGGCGGCGTACCTCTATCGCGGAGACGTC 1540
 DB 356 AGGCGGCGGTGCTGTAATTTAAACGAGAACTTCGACCAAGCTGAAGCGAGATG 415
 QY 1541 TTCCTACCGCGGCGGAGCCGAGCGATCGAGTCATATCCCGTGTGCGCCAGACC 1600
 DB 416 TATATACCGGAGATGTAACCAAGCCATGAGATGATGATTTCTTCCGGAAT 475
 QY 1601 GCCGCGCGCAATCTGCTCCCGAGCGGCTACCCAACTACGAGGCGCGCGCG 1660
 DB 476 CCATCGCGCAATCTTACTTCAAGCGCGGATATCTCACTAGATGCTGTCTGTC 535
 QY 1661 TTCAACAGGCTGAGGTCGCGCATTTGACCTCATCCCGAGCAAGGGGTGGAGATCGAC 1720
 DB 536 TATAGCGGCTGAGATTCGCAATACGATCTTCTCCGAGAGATGATGGAAATCAAT 595
 QY 1721 ATCGACTGCTGGAATTCATGCGCGCAAGAACACACCGCATATGTCATTAACCC 1780
 DB 596 CTCGATGCGCTGAGAGCGGCTGCGGATGAGATACCGTCGCAATGATATCAACCC 655
 QY 1781 AACACCGCGCGGAGCGCTTACTCTTACGACCATCTGTGCCAGGT 1827
 DB 656 AACATTCATGAGAAAGCTTACACCTACGACCATCTCAACAGGT 702

RESULT 11

ADA70161
 ID ADA70161 standard; DNA; 930 BP.

AC ADA70161;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 3484.

XX plant; bacterial infection; fungal infection; viral infection; rice;
 KM gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

PF 22-JUN-2001; 2001WO-1B001105.

XX 22-JUN-2001; 2001WO-1B001105.

XX (SYN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
 PI Katagiri F, Quan S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;

DR WPI; 2003-175390/17.

XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

PS Claim 6; SEQ ID NO 3484; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes

CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC the incompatible interaction of plant gene expression in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 930 BP; 217 A; 259 C; 229 G; 224 T; 0 U; 1 Other;

Query Match 1.3%; Score 141.4; DB 8; Length 930;
 Best Local Similarity 64.1%; Pred. No. 4e-14;
 Matches 230; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 1478 CACAGCGCGGAGAGACCTGTGCGAGGCGGCGTACCTATCGCGCGAGAC 1537
 DB 274 CGCGAGCTATTCAGAGTACTTATCTGATCTTCTTCAAGCTTTCAGACATGAT 333
 QY 1538 GTCTTCTCACCGCGCGGAGCCAGCGCATCGAGTCATATCCCGTGTGCGCCAG 1597
 DB 334 ATTTCTTCACATCTGAGAGTACCCAGCAATCGAGATGTTATGTCTTTTGGCCAA 393
 QY 1598 ACCGCGCGGCAATCTGCTCCCGAGCGGCGGCTACCCAACTAGAGGCGCGCC 1657
 DB 394 ---CGAGTGCATATATTTGCTTTCAGAGCGCGGTACCAAGAGAGACATGCG 450
 QY 1658 GCGTTCAACAGGCTGAGGTCGCGCATTTGACCTCATCCCGCAAGGGGTGGAGATC 1717
 DB 451 GTGTTCCAGAGATGAGAGTGGGCTCATATCTTGTTCAGAGAGAGATGGAGATT 510
 QY 1718 GACATGACTGCTGGAATTCATGCGCGCAAGAACACACCGCATGTCATCAAAAC 1777
 DB 511 AATGTGAAGCTGTGAAGCTTTAGCAGATGAGATACTGTTCGAATGATGATTAATA 570
 QY 1778 CCCAACACCGGCGGAGGCTTACTCTACGACCATCTGTCCAAAGTTTACATCC 1836
 DB 571 CCCATTAACCTTGTGTATGTTGACCTTATGAGCATCTGTCCAAATTCAGATAC 629

RESULT 12

ADJ40562
 ID ADJ40562 standard; cDNA; 930 BP.

AC ADJ40562;

DT 06-MAY-2004 (first entry)

DE Plant cDNA #1562.

XX plant; gene; ss; transcription; plant genome augmentation; cereal;
 KM soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
 KM maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
 KM stress tolerance; salt tolerance; cold tolerance; drought tolerance;
 KM plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
 KM antifungal.

XX Eukaryota.

XX US2004016025-A1.

XX 22-JAN-2004.

PF 26-SEP-2002; 2002US-00260238.

XX 26-SEP-2001; 2001US-0325277P.

PR 26-SEP-2001; 2001US-0325448P.

PR 04-APR-2002; 2002US-0370620P.

XX (BUDW/) BUDWORTH P.

PA (MOUG/) MOUGAMER T.

PA (BRIG/) BRIGGS S P.

PA (COOP/) COOPER B.
 PA (GLAZ/) GLAZEBROOK J.
 PA (GOFF/) GOFF S. A.
 PA (KATRI/) KATRIGRI F.
 PA (KREP/) KREPS J.
 PA (PROV/) PROVART N.
 PA (RICK/) RICHE D.
 PA (ZHUT/) ZHU T.
 XX Budworth P., Moughamer T., Briggs SP., Cooper B., Glazebrook J;
 PI Goff SA, Katrigiri F, Kreps J, Provart N, Riche D, Zhu T;
 XX WPI; 2004-190374/18.
 DR
 XX
 PT New rice promoter, useful for manipulating crop plants to alter or
 PT improve phenotypic characteristics, e.g. produce large quantities of oil
 PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
 PT or high nutritional value.
 PS
 XX
 PS Claim 25; SEQ ID NO 1562; 230pp; English.
 CC The invention relates to plant nucleotide sequences that direct seed-
 CC leaf-and/or stem-, panicle-, root- or pollen-specific or -preferential
 CC or constitutive transcription of an operatively linked nucleic acid
 CC segment. The invention also relates to a method for augmenting a plant
 CC genome and a method of identifying a gene, where its expression is
 CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
 CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
 CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
 CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
 CC encode are useful for manipulating crop plants to alter or improve
 CC phenotypic characteristics, to produce large quantities of oil or
 CC proteins, to incur resistance to insecticides, viruses or fungi, and to
 CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
 CC have a high nutritional value with reduced apical dominance or dwarfism,
 CC early flowering or altered metabolic pathways. This sequence represents a
 CC plant nucleic acid of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification but was obtained in
 CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 930 BP; 217 A; 259 C; 229 G; 224 T; 0 U; 1 Other;
 Query Match 1.3%; Score 141.4; DB 12; Length 930;
 Best Local Similarity 64.1%; Pred. No. 4e-14;
 Matches 230; Conservative 0; Mismatches 126; Indels 3; Gaps 1;
 QY 1478 CACAGCGCGGCGGAGACACCTGTCGAGGCGCGCTACATGCTATCGGCCGAC 1537
 DB 274 CGCCGAGCTATTTGACAGATACCTATCTGTGATCTTCTTACAGCTTTGACAGATGAT 333
 QY 1538 GTCTTCTTACCGCGCGGAGACCGAGCGATCGAGTCATATATCCCGTGTGCGCCAG 1597
 DB 334 ATTTTCTTCACTCTGAGAGTACCCAGCAATCGAGATTTGATATCTGTTTGGCAA 393
 QY 1598 ACCGCGGCGCCACATTTCTGCTCCCAAGCCAGGCTACCCAACTACGAGCGCGCC 1657
 DB 394 ---CGAGGTGCAATATATGCTTCCAAAGCCGGGTACCCAAACATGAGACATGCG 450
 QY 1658 GCGTTCAACAGGCTGAGAGTCCGCGATTTGCACTTCATCCCGCAAGGGGTGGAGATC 1717
 DB 451 GTGTTCACAGAGATGAAAGTGGGCTCTATGATCTTGTTCAGAGAGAGATGGAGATT 510
 QY 1718 GACATCGACTGCTGGAATTCATCGCCGACAGAACACACCGCGATGATCTCATTAAC 1777
 DB 511 AATGTGAAGCTGTGTAAGCTTTAGCAATGAGAAATATCTGTGCAATGATGATTAAC 570
 QY 1778 CCGAACAACCGGTGCGGAGCGTTTACTTCTTACGACATCTGTCCAGGTTTCATATCC 1836
 DB 571 CCCAATAACCCCTTGTGTATGTAAGTATGAGCACTTGTGAGCACTGTCCAAAGATGCAATAC 629

ID AAD57651 standard; DNA; 1200 BP.
 XX
 XX AAD57651;
 AC
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Rice enhanced yield gene, NMA1.
 XX
 XX Rice; abiotic stress tolerance; pathogen resistance; disease resistance;
 KM grain quality; nutritional content; plant yield; NMA1; plant; gene; ds.
 XX
 OS Oryza sativa.
 XX
 FH Key Location/Qualifiers
 FH CDS 1..1200
 FT /*tag= a
 FT /product= "Rice NMA1 protein"
 XX
 XX MO2003048319-A2.
 XX
 PD 12-JUN-2003.
 XX
 XX 27-NOV-2002; 2002MO-US038359.
 XX
 XX 30-NOV-2001; 2001US-0334501P.
 XX
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA Sainz MB, Salmeron J, Weislo L;
 PI WPI; 2003-505288/47.
 DR P-PSDB; AAE38283.
 XX
 XX New nucleic acid from Oryza sativa, useful for altering abiotic stress
 PT tolerance, pathogen or disease resistance or the grain quality,
 PT nutritional content or yield in a plant.
 XX
 XX Claim 2; Page 202-203; 223pp; English.
 CC The invention relates to nucleic acid molecules from rice encoding
 CC proteins for abiotic stress tolerance, enhanced pathogen or disease
 CC resistance and altered nutritional quality. The sequences of the
 CC invention are useful for altering abiotic stress tolerance, pathogen or
 CC disease resistance or the grain quality, nutritional content or yield in
 CC a plant. The present sequence is rice enhanced yield gene, NMA1
 CC
 XX
 SQ Sequence 1200 BP; 272 A; 282 C; 354 G; 292 T; 0 U; 0 Other;
 Query Match 1.3%; Score 139.6; DB 10; Length 1200;
 Best Local Similarity 64.1%; Pred. No. 8.6e-14;
 Matches 227; Conservative 0; Mismatches 124; Indels 3; Gaps 1;
 QY 1477 ACAAGGCGCGTGACAGACACCTGTGCGAGGCGGTCCGATACATGCTATCGCGACGA 1536
 DB 186 AGGAGGTATATCGCGGATCTTATCGAGACTTCCATATAGCTATCACTGATGA 245
 QY 1537 GCTTCTTCAACCGCGCGGAGACCGAGCGATCGAGTCATATATCCCGTGTGCGCCA 1596
 DB 246 TGTGTACTGACAGATGGCTGTGTCTCAAGGATGAGATATCTGCTCTAGCTCG 305
 QY 1597 GACCGCGCGCCCAACTTCTGCTCCCAAGCCAGGCTACCCAAATACAGAGCGCGC 1656
 DB 306 ---CCCTGTGTCCAACATCTCTGTGCCCAAGCGGTACTCTGTTCACAGGAGACCGGC 362
 QY 1657 CGGTTCAACAGGCTGAGAGTCCGCGATTTGACCTCATCCCGACAGAGGGTGGAGAT 1716
 DB 363 AGTTTCAATGAGATGAGAGGTCAAGTACTTTGATCTTCTCCAGAGAGTGTGGAGAT 422
 QY 1717 CGACATCGACTCGTGTGAATTCATCGCGACAGAACACACCGCCATGTGTCATATAA 1776
 DB 423 TGATCTTGATGAGATGACAGAACTTGCTGACAAAGAACAGGTTTCAATGTCATATCAA 482
 QY 1777 CCGCAACAACCGGTGCGGAGCGTTTACTCTTACGACATCTGTCCAAAGTTTC 1830

Db 483 TCAGAGAAATCCCTGTGGCAATGTACACTTGTGAGCAATTTGGCAAGGTTGC 536

RESULT 14

ID ADA70274 standard; DNA; 1290 BP.

ADA70274;

AC ADA70274;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 3597.

XX Plant, bacterial infection; fungal infection; viral infection; rice;

XX gene; db.

OS Oryza sativa.

PN WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-1B001105.

XX 22-JUN-2001; 2001WO-1B001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,

XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G,

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to

XX pathogenic infection for conferring resistance or tolerance to a plant to

XX bacterial, fungal or viral infection by determining or detecting plant

XX gene expression.

XX Claim 6; SEQ ID NO 3597; 899bp; English.

XX The present invention relates to a method (M1) for identifying genes

XX involved in plant resistance or response to pathogenic infection. M1

XX comprises identifying a gene whose expression is significantly altered in

XX the incompatible interaction of plant gene expression relative to

XX expression of the gene in an uninfected plant, in a mutant plant that

XX does not express a gene associated with response to pathogenic infection,

XX or in a corresponding incompatible or compatible interaction. (M1) is

XX useful for conferring resistance to resistance or tolerance to a plant to

XX bacterial, fungal or viral infection. The present sequence was used to

XX illustrate the invention.

XX Sequence 1290 BP; 280 A; 307 C; 400 G; 297 T; 0 U; 6 Other;

XX Query Match 1.3%; Score 137.4; DB 8; Length 1290;

XX Best Local Similarity 60.1%; Pred. No. 2.1e-13;

XX Matches 247; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

XX

OY 7202 CGGCAATTCAAGAGTGGATCCGCACTTGACCTCATCCGCCGCAAGGGGTGGAGA 7261

Db 491 CAGGTGTTCAATGAGCATGAGGTCAGTACTTTGATCTTCCAGAGATGCTGGAGG 550

OY 7262 TCGACATGACATCGCTGGAATTCATCGCCGACAGAACCAACCGGATGATCATATA 7321

Db 551 TTGATCTTGATGAGAGTGGAGCACTTGTGACAAAGAACGGTTGGCATGTGATATCA 610

OY 7322 ACCCAACCAATCCGTGCGGACAGCTTACTCTAGACCATGTGGCCAGG 7372

Db 611 ATCCAGAAATCCCTGTGGCAATGTACACTTGTGAGCATTTGGCCAGG 661

RESULT 15

ID ABZ14179 standard; DNA; 1389 BP.

ABZ14179;

XX 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1984.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-026647P.

XX 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES. INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Krens J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and

XX producing plants with increased tolerance to these abiotic stresses.

XX Claim 14; SEQ ID NO 1984; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant

XX cell has been exposed, comprising: (a) contacting nucleic acid

XX representative of expressed polynucleotides in the plant cell with an

XX array or probes representative of the plant cell genome; and (b)

XX detecting a profile of expressed polynucleotides in the plant cell

XX characteristic of a stress response. The method is useful in the

XX production of transgenic plants, cells and seeds and in producing plants

XX with increased tolerance to abiotic stress. The present sequence is that

XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used

XX in methods of the invention. Note: The sequence data for this patent is

XX not represented in the printed specification but is based on sequence

XX information supplied to Derwent by the European Patent Office

XX Sequence 1389 BP; 389 A; 322 C; 337 G; 341 T; 0 U; 0 Other;

XX Query Match 1.2%; Score 136; DB 6; Length 1389;

XX Best Local Similarity 63.6%; Pred. No. 3.6e-13;

XX Matches 224; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

XX

OY 1482 GCGCGTGGCAGAGCACTGTGCGAGGGGCGTCCGTAATGTTATCGCGGACGAGCTT 1541

Db 344 GAGCGGTGCTGATTAATTAAGAACGAGATCTTCCGCAAGATTAACGCTGAAGATATTT 403


```
PF 04-JUL-2000; 2000WO-JP004425.
XX
XX 05-JUL-1999; 99JP-00190318.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Mori S, Nakaniishi H, Takahashi M, Nishizawa N,
XX WPI; 2001-138030/14.
XX
XX Gramineous plant, e.g. rice, with tolerance to iron deficiency for growth
XX in calcareous alkaline soil is constructed by transformation with a gene
XX of encoding an enzyme of the mugineic acids biosynthetic pathway.
XX
XX Example 6; Page 20; 61pp; Japanese.
XX
XX The present invention describes a method for constructing a rice plant
XX with improved iron absorbability and a tolerance to iron deficiency which
XX comprises transferring a gene encoding an enzyme in the mugineic acid
XX biosynthetic pathway into a rice plant. The method is for constructing
XX gramineous plant e.g. rice with tolerance to iron deficiency, which is
XX useful in agriculture in producing new breeds of rice plants capable of
XX vigorous growth in calcareous alkaline soil for improving crop
XX production. The constructed plant has tolerance to iron deficiency, and
XX is therefore capable of vigorous growth in calcareous alkaline soil. The
XX present sequence represents a primer which is used in an example from the
XX
XX Sequence 24 BP; 6 A; 7 C; 5 G; 6 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 24; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 TCTGGATCTCAGTCATCCACA 24
DB 1 TCTGGATCTCAGTCATCCACA 24
RESULT 2
AAFP32380
ID AAFP32380 standard; cDNA; 10966 BP.
XX
XX AAFP32380;
AC
XX 18-APR-2001 (first entry)
XX
XX Hordeum vulgare L. var. Igr1 NAAT encoding cDNA SEQ ID NO:1.
XX
XX Hordeum vulgare L. var. Igr1; nicotianamine aminotransferase; NAAT;
XX NAAT-A; NAAT-B; iron deficiency; gramineous plant; barley; rice;
XX mugineic acid biosynthetic pathway; calcareous alkaline soil; ss.
XX
XX Hordeum vulgare.
OS
XX WO200101762-A1.
XX
XX 11-JAN-2001.
XX
XX 04-JUL-2000; 2000WO-JP004425.
XX
XX 05-JUL-1999; 99JP-00190318.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Mori S, Nakaniishi H, Takahashi M, Nishizawa N;
XX WPI; 2001-138030/14.
XX
XX P-PSDB; AAB69048. AAB69049.
XX
XX Gramineous plant, e.g. rice, with tolerance to iron deficiency for growth
XX in calcareous alkaline soil is constructed by transformation with a gene
XX of encoding an enzyme of the mugineic acids biosynthetic pathway.
```

```
XX
XX Claim 6; Fig 10; 61pp; Japanese.
XX
XX The present invention describes a method for constructing a rice plant
XX with improved iron absorbability and a tolerance to iron deficiency which
XX comprises transferring a gene encoding an enzyme in the mugineic acid
XX biosynthetic pathway into a rice plant. The method is for constructing
XX gramineous plant e.g. rice with tolerance to iron deficiency, which is
XX useful in agriculture in producing new breeds of rice plants capable of
XX vigorous growth in calcareous alkaline soil for improving crop
XX production. The constructed plant has tolerance to iron deficiency, and
XX is therefore capable of vigorous growth in calcareous alkaline soil. The
XX present sequence encodes two nicotianamine aminotransferases (NAAT),
XX designated NAAT-A and NAAT-B, isolated from Hordeum vulgare L. var. Igr1
XX (barley), for use in the method of the invention
XX
XX Sequence 10966 BP; 3082 A; 2422 C; 2366 G; 3096 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 24; DB 5; Length 10966;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 TCTGGATCTCAGTCATCCACA 24
DB 8475 TCTGGATCTCAGTCATCCACA 8498
RESULT 3
AAS27894/C
ID AAS27894 standard; cDNA; 104 BP.
XX
XX AAS27894;
AC
XX 07-NOV-2001 (first entry)
XX
XX Novel cDNA encoding for human respiratory antigen #26.
XX
XX Human; respiratory antigen; respiratory disorder; throat disorder;
XX lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
XX anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX respiratory active; ss.
XX
XX Homo sapiens.
OS
XX WO200155448-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001333.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 11-JUL-2000; 2000US-0217496P.
XX
XX 14-JUL-2000; 2000US-0218290P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
```

PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227109P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233397P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249256P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0251997P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-476224/51.
XX P-PSDB; AAU17710.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the respiratory system including respiratory cancers
XX and also for testing and detection e.g. diagnosis.
XX
XX Claim 4; SEQ ID No 36; 546bp; English.
XX
XX The present invention relates to the isolation of novel human respiratory
XX antigens (AAU17685-AAU17915), and cDNA and genomic sequences encoding for
XX these polypeptides. The sequences of the invention are useful for
XX preventing, treating and/or prognosing disorders related to the
XX respiratory system including throat disorders (e.g. vocal cord paralysis,
XX tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
XX disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
XX disorders and cancers of the respiratory tissues e.g. lung cancer. The
XX polynucleotide sequences of the invention are useful in gene therapy and
XX antisense therapy. AAS27869-AAS28159 encode for novel human respiratory
XX antigens. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 104 BP; 32 A; 19 C; 23 G; 29 T; 0 U; 1 Other;

Query Match 80.0%; Score 19.2; DB 4; Length 104;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TCTGGATCTCAGTCATCCACA 24
 DB 68 TCTGGCTTTCAGTCATCCATCA 45

RESULT 4
 ADG40798/C
 ID ADG40798 standard; cDNA, 104 BP.

AC ADG40798;

DT 26-FEB-2004 (fixed entry)

DE Human respiratory system associated protein cDNA seq id 36.

XX antiinflammatory; antiallergic; antiaesthetic; cytostatic; gene therapy;
 XX respiratory system antigen;
 KW human respiratory system associated polynucleotide;
 KW human respiratory system associated polynucleotide;
 KW respiratory system disorder; throat disorder; vocal cord paralysis;
 KW tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
 KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
 KW histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
 KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
 KW cancer of the nose; gene therapy; chromosome identification; forensic;
 KW human respiratory system associated protein; gene; ss; human.

OS Homo sapiens.

PN US2003215893-A1.

XX 20-NOV-2003.

PF 07-AUG-2002; 2002US-00212872.

XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 09-MAY-2000; 2000US-0205515P.
 PR 17-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225575P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 30-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 28-SEP-2000; 2000US-0235935P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240606P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-024474P.
 PR 08-NOV-2000; 2000US-024475P.
 PR 08-NOV-2000; 2000US-024476P.
 PR 08-NOV-2000; 2000US-024477P.
 PR 08-NOV-2000; 2000US-024478P.
 PR 08-NOV-2000; 2000US-024523P.
 PR 08-NOV-2000; 2000US-024524P.
 PR 08-NOV-2000; 2000US-024525P.
 PR 08-NOV-2000; 2000US-024526P.
 PR 08-NOV-2000; 2000US-024527P.
 PR 08-NOV-2000; 2000US-024528P.
 PR 08-NOV-2000; 2000US-024532P.
 PR 08-NOV-2000; 2000US-024533P.
 PR 08-NOV-2000; 2000US-024609P.
 PR 08-NOV-2000; 2000US-024610P.
 PR 08-NOV-2000; 2000US-024611P.
 PR 17-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-0076466O.
PR 14-FEB-2002; 2002US-0007409S.

PA (HUMA-) HUMAN GENOME SCT INC.
XX
XX
XX
PI Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-902033/82.
DR MPI; 2003-902033/82.
DR P-PDB; ADG41090.

XX
FT Novel respiratory system antigen and polynucleotides encoding the
PT polypeptides, useful for treating diagnosing, treating or preventing
PT tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
PT cancer.

PS Claim 1; SEQ ID NO 36; 236pp; English.

XX
XX The invention describes an isolated polypeptide (I) comprising an amino acid sequence that is at least 90% identical to polypeptide fragment of any one of 239 respiratory system antigen sequences (PS) and having biological activity, polypeptide domain or epitope of PS, full-length protein of PS, or variant, allelic variant or species homolog of PS, (I) or a polynucleotide (II) encoding (I) is also useful for diagnosing a pathological condition or a susceptibility to a pathological condition in a subject which involves determining the presence or absence of mutation in (II) or determining the presence or amount of expression of (I) in a biological sample and diagnosing a pathological condition based on the result. The human respiratory system associated polynucleotides, the polypeptides encoded by them, and antibodies that immunospecifically bind these polypeptides are useful in diagnosis, treatment, prevention and/or prognosis of disorders of respiratory system such as throat disorders (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis, sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and CC cancer of the nose). The polynucleotides are useful in gene therapy

Query Match	80.0%;	Score 19.2;	DB 10;	Length 104;
Best Local Similarity	87.5%;	Pred. No. 23;		
Matches 21; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy 1 TCTGGATCTCAGTGCATCCACA 24
| | | | | | | | | |
Db 68 TCTGGGCTTCAGTGATCCATCA 45

RESULT 5

AD196572/c	ID	AD196572 standard; DNA; 104 BP.
AD196572;	AC	
AD196572;	XX	
04-NOV-2004 (first entry)	DT	
	XX	
Human respiratory system associated gene SegID36.	DE	
respiratory system-related polypeptide; antiasthmatic; antibacterial;	KW	
pneumonia; cytoskeletal; antiasthmatic; antiasthmatic; gene therapy;	KW	
pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;	KW	
anemia; leukemia; inflammation; sinusitis;	KW	
chronic obstructive pulmonary disease; infectious disease; human; gene	KW	
ds.	KW	
Homo sapiens.	OS	
US200307704-A1.	PN	
24-APR-2003.	PD	
	XX	
14-FEB-2002; 2002US-00074095.	PF	
	XX	
31-JAN-2000; 2000US-0179065P.	PR	
04-FEB-2000; 2000US-0180628P.	PR	
24-FEB-2000; 2000US-0184664P.	PR	
02-MAR-2000; 2000US-0186350P.	PR	
16-MAR-2000; 2000US-0198874P.	PR	
17-MAR-2000; 2000US-0199076P.	PR	
18-APR-2000; 2000US-0198123P.	PR	
19-MAY-2000; 2000US-0205513P.	PR	
07-JUN-2000; 2000US-0209467P.	PR	
28-JUN-2000; 2000US-0214886P.	PR	
30-JUN-2000; 2000US-0215135P.	PR	
07-JUL-2000; 2000US-0216647P.	PR	
07-JUL-2000; 2000US-0216880P.	PR	
11-JUL-2000; 2000US-0217487P.	PR	
11-JUL-2000; 2000US-0217496P.	PR	
14-JUL-2000; 2000US-0218290P.	PR	
26-JUL-2000; 2000US-0220963P.	PR	
26-JUL-2000; 2000US-0220964P.	PR	
14-AUG-2000; 2000US-0224518P.	PR	
14-AUG-2000; 2000US-0224519P.	PR	
14-AUG-2000; 2000US-0225213P.	PR	
14-AUG-2000; 2000US-0225214P.	PR	
14-AUG-2000; 2000US-0225214P.	PR	
14-AUG-2000; 2000US-0225266P.	PR	
14-AUG-2000; 2000US-0225266P.	PR	
14-AUG-2000; 2000US-0225270P.	PR	
14-AUG-2000; 2000US-0225471P.	PR	
14-AUG-2000; 2000US-0225757P.	PR	
14-AUG-2000; 2000US-0225758P.	PR	
14-AUG-2000; 2000US-0225759P.	PR	
18-AUG-2000; 2000US-0226681P.	PR	
22-AUG-2000; 2000US-0226681P.	PR	
22-AUG-2000; 2000US-0226868P.	PR	
22-AUG-2000; 2000US-0227182P.	PR	
23-AUG-2000; 2000US-0227009P.	PR	
30-AUG-2000; 2000US-0228924P.	PR	
01-SEP-2000; 2000US-0229287P.	PR	
01-SEP-2000; 2000US-0229343P.	PR	
01-SEP-2000; 2000US-0229344P.	PR	
01-SEP-2000; 2000US-0229345P.	PR	
05-SEP-2000; 2000US-0229509P.	PR	
05-SEP-2000; 2000US-0229513P.	PR	
06-SEP-2000; 2000US-0230437P.	PR	
06-SEP-2000; 2000US-0230438P.	PR	
08-SEP-2000; 2000US-0231242P.	PR	
08-SEP-2000; 2000US-0231243P.	PR	
08-SEP-2000; 2000US-0231244P.	PR	
08-SEP-2000; 2000US-0231413P.	PR	
08-SEP-2000; 2000US-0231414P.	PR	


```
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249267P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0293900P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251566P.
PR 06-DEC-2000; 2000US-0251568P.
PR 06-DEC-2000; 2000US-0251689P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251900P.
PR 11-DEC-2000; 2000US-0251907P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764860.

XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX PI Rosen CA, Ruben SM, Barash SC;
XX
XX DR WPI; 2003-765403/72.
XX DR E-PSDB; ADI96864.
XX
XX PT New human respiratory system-related polypeptide and genes, useful for
XX PT treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic
XX PT fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or
XX PT sinusitis.
XX
XX PS Claim 1; SEQ ID NO 36; 202pp; English.
XX
XX CC This invention is related to a novel isolated polypeptide, which
XX CC comprises a human respiratory system-related polypeptide, and the DNA
XX CC sequence which encodes it. The invention may be useful for the
XX CC development of compounds with an antiasthmatic, antibacterial,
XX CC antiinflammatory, cytostatic, antianaemic or antiallergic activity. In
XX CC addition, the sequences disclosed may be useful for gene therapy. The
XX CC polypeptide or polynucleotide is useful for treating, preventing or
XX CC ameliorating a medical condition, for example pneumonia, lung cancer,
XX CC cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukemia,
XX CC inflammations, sinusitis, chronic obstructive pulmonary disease or
XX CC infectious diseases. The polypeptide or polynucleotide is also useful for
XX CC diagnosing any of these diseases or a susceptibility to the disease. The
XX CC present sequence is that of a respiratory system associated human gene of
XX CC the invention.
XX
XX SQ Sequence 104 BP; 32 A; 19 C; 23 G; 29 T; 0 U; 1 Other;

Query Match 80.0%; Score 19.2; DB 11; Length 104;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCTGGGATCTCAGTCATCCACAA 24
DB 68 TCTGGGCTTCAGTCATCCATCA 45

RESULT 6
AAS28222/c
ID AAS28222 standard; DNA; 224 BP.
XX
XX AC AAS28222;
XX
XX DT 07-NOV-2001 (first entry)
XX
XX DE Genomic sequence #62 encoding for novel human respiratory antigen.
XX
XX KW Human; respiratory antigen; respiratory disorder; throat disorder;
XX KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
XX KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
XX KW respiratory active; ds.
XX
XX OS Homo sapiens.
XX
```

PN W0200155448-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001333.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186550P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225577P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231342P.
PR 08-SEP-2000; 2000US-0231433P.
PR 08-SEP-2000; 2000US-0231444P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246603P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX
PI

XX WPI, 2001-476224/51.
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the respiratory system including respiratory cancers
PT and also for testing and detection e.g. diagnosis.
XX
PS Disclosure, SEQ ID No 656, 546bp, English.
XX
XX The present invention relates to the isolation of novel human respiratory
CC antigens (AAU17685-AAU17975), and cDNA and genomic sequences encoding for
CC these polypeptides. The sequences of the invention are useful for
CC preventing, treating and/or prognosing disorders related to the
CC respiratory system including throat disorders (e.g. vocal cord paralysis,
CC tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
CC disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
CC disorders and cancers of the respiratory tissues e.g. lung cancer. The
CC polynucleotide sequences of the invention are useful in gene therapy and
CC antisense therapy. AAS28161-AAS28764 represent genomic sequences encoding
CC for novel human respiratory antigens. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 224 BP; 76 A; 39 C; 58 G; 51 T; 0 U; 0 Other;
Query Match 80.0%; Score 19.2; DB 4; Length 224;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 1 TCTGGATCTCAGTCATCCACA 24
Db 162 TCTGGCTTTCAGTCATCCATCA 139
RESULT 7
AAS28221/c
ID AAS28221 standard; DNA, 224 BP.
XX
AC AAS28221;
XX
DT 07-NOV-2001 (first entry)
XX
XX Genomic sequence #61 encoding for novel human respiratory antigen.
DE
XX Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW respiratory active; ds.
XX
XX Homo sapiens.
OS
XX
XX MO200155448-A1.
PN
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001333.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 19-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229345P.
PR 01-SEP-2000; 2000US-0229349P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236379P.
PR 29-SEP-2000; 2000US-0236380P.
PR 29-SEP-2000; 2000US-0236381P.
PR 29-SEP-2000; 2000US-0236382P.
PR 29-SEP-2000; 2000US-0236383P.
PR 29-SEP-2000; 2000US-0236384P.
PR 29-SEP-2000; 2000US-0236385P.
PR 29-SEP-2000; 2000US-0236386P.
PR 29-SEP-2000; 2000US-0236387P.
PR 29-SEP-2000; 2000US-0236388P.
PR 29-SEP-2000; 2000US-0236389P.
PR 29-SEP-2000; 2000US-0236390P.
PR 29-SEP-2000; 2000US-0236391P.
PR 29-SEP-2000; 2000US-0236392P.
PR 29-SEP-2000; 2000US-0236393P.
PR 29-SEP-2000; 2000US-0236394P.
PR 29-SEP-2000; 2000US-0236395P.
PR 29-SEP-2000; 2000US-0236396P.
PR 29-SEP-2000; 2000US-0236397P.
PR 29-SEP-2000; 2000US-0236398P.
PR 29-SEP-2000; 2000US-0236399P.
PR 29-SEP-2000; 2000US-0236400P.
PR 29-SEP-2000; 2000US-0236401P.
PR 29-SEP-2000; 2000US-0236402P.
PR 29-SEP-2000; 2000US-0236403P.
PR 29-SEP-2000; 2000US-0236404P.
PR 29-SEP-2000; 2000US-0236405P.
PR 29-SEP-2000; 2000US-0236406P.
PR 29-SEP-2000; 2000US-0236407P.
PR 29-SEP-2000; 2000US-0236408P.
PR 29-SEP-2000; 2000US-0236409P.
PR 29-SEP-2000; 2000US-0236410P.
PR 29-SEP-2000; 2000US-0236411P.
PR 29-SEP-2000; 2000US-0236412P.
PR 29-SEP-2000; 2000US-0236413P.
PR 29-SEP-2000; 2000US-0236414P.
PR 29-SEP-2000; 2000US-0236415P.
PR 29-SEP-2000; 2000US-0236416P.
PR 29-SEP-2000; 2000US-0236417P.
PR 29-SEP-2000; 2000US-0236418P.
PR 29-SEP-2000; 2000US-0236419P.
PR 29-SEP-2000; 2000US-0236420P.
PR 29-SEP-2000; 2000US-0236421P.
PR 29-SEP-2000; 2000US-0236422P.
PR 29-SEP-2000; 2000US-0236423P.
PR 29-SEP-2000; 2000US-0236424P.
PR 29-SEP-2000; 2000US-0236425P.
PR 29-SEP-2000; 2000US-0236426P.
PR 29-SEP-2000; 2000US-0236427P.
PR 29-SEP-2000; 2000US-0236428P.
PR 29-SEP-2000; 2000US-0236429P.
PR 29-SEP-2000; 2000US-0236430P.
PR 29-SEP-2000; 2000US-0236431P.
PR 29-SEP-2000; 2000US-0236432P.
PR 29-SEP-2000; 2000US-0236433P.
PR 29-SEP-2000; 2000US-0236434P.
PR 29-SEP-2000; 2000US-0236435P.
PR 29-SEP-2000; 2000US-0236436P.
PR 29-SEP-2000; 2000US-0236437P.
PR 29-SEP-2000; 2000US-0236438P.
PR 29-SEP-2000; 2000US-0236439P.
PR 29-SEP-2000; 2000US-0236440P.
PR 29-SEP-2000; 2000US-0236441P.
PR 29-SEP-2000; 2000US-0236442P.
PR 29-SEP-2000; 2000US-0236443P.
PR 29-SEP-2000; 2000US-0236444P.
PR 29-SEP-2000; 2000US-0236445P.
PR 29-SEP-2000; 2000US-0236446P.
PR 29-SEP-2000; 2000US-0236447P.
PR 29-SEP-2000; 2000US-0236448P.
PR 29-SEP-2000; 2000US-0236449P.
PR 29-SEP-2000; 2000US-0236450P.
PR 29-SEP-2000; 2000US-0236451P.
PR 29-SEP-2000; 2000US-0236452P.
PR 29-SEP-2000; 2000US-0236453P.
PR 29-SEP-2000; 2000US-0236454P.
PR 29-SEP-2000; 2000US-0236455P.
PR 29-SEP-2000; 2000US-0236456P.
PR 29-SEP-2000; 2000US-0236457P.
PR 29-SEP-2000; 2000US-0236458P.
PR 29-SEP-2000; 2000US-0236459P.
PR 29-SEP-2000; 2000US-0236460P.
PR 29-SEP-2000; 2000US-0236461P.
PR 29-SEP-2000; 2000US-0236462P.
PR 29-SEP-2000; 2000US-0236463P.
PR 29-SEP-2000; 2000US-0236464P.
PR 29-SEP-2000; 2000US-0236465P.
PR 29-SEP-2000; 2000US-0236466P.
PR 29-SEP-2000; 2000US-0236467P.
PR 29-SEP-2000; 2000US-0236468P.
PR 29-SEP-2000; 2000US-0236469P.
PR 29-SEP-2000; 2000US-0236470P.
PR 29-SEP-2000; 2000US-0236471P.
PR 29-SEP-2000; 2000US-0236472P.
PR 29-SEP-2000; 2000US-0236473P.
PR 29-SEP-2000; 2000US-0236474P.
PR 29-SEP-2000; 2000US-0236475P.
PR 29-SEP-2000; 2000US-0236476P.
PR 29-SEP-2000; 2000US-0236477P.
PR 29-SEP-2000; 2000US-0236478P.
PR 29-SEP-2000; 2000US-0236479P.
PR 29-SEP-2000; 2000US-0236480P.
PR 29-SEP-2000; 2000US-0236481P.
PR 29-SEP-2000; 2000US-0236482P.
PR 29-SEP-2000; 2000US-0236483P.
PR 29-SEP-2000; 2000US-0236484P.
PR 29-SEP-2000; 2000US-0236485P.
PR 29-SEP-2000; 2000US-0236486P.
PR 29-SEP-2000; 2000US-0236487P.
PR 29-SEP-2000; 2000US-0236488P.
PR 29-SEP-2000; 2000US-0236489P.
PR 29-SEP-2000; 2000US-0236490P.
PR 29-SEP-2000; 2000US-0236491P.
PR 29-SEP-2000; 2000US-0236492P.
PR 29-SEP-2000; 2000US-0236493P.
PR 29-SEP-2000; 2000US-0236494P.
PR 29-SEP-2000; 2000US-0236495P.
PR 29-SEP-2000; 2000US-0236496P.
PR 29-SEP-2000; 2000US-0236497P.
PR 29-SEP-2000; 2000US-0236498P.
PR 29-SEP-2000; 2000US-0236499P.
PR 29-SEP-2000; 2000US-0236500P.
PR 29-SEP-2000; 2000US-0236501P.
PR 29-SEP-2000; 2000US-0236502P.
PR 29-SEP-2000; 2000US-0236503P.
PR 29-SEP-2000; 2000US-0236504P.
PR 29-SEP-2000; 2000US-0236505P.
PR 29-SEP-2000; 2000US-0236506P.
PR 29-SEP-2000; 2000US-0236507P.
PR 29-SEP-2000; 2000US-0236508P.
PR 29-SEP-2000; 2000US-0236509P.
PR 29-SEP-2000; 2000US-0236510P.
PR 29-SEP-2000; 2000US-0236511P.
PR 29-SEP-2000; 2000US-0236512P.
PR 29-SEP-2000; 2000US-0236513P.
PR 29-SEP-2000; 2000US-0236514P.
PR 29-SEP-2000; 2000US-0236515P.
PR 29-SEP-2000; 2000US-0236516P.
PR 29-SEP-2000; 2000US-0236517P.
PR 29-SEP-2000; 2000US-0236518P.
PR 29-SEP-2000; 2000US-0236519P.
PR 29-SEP-2000; 2000US-0236520P.
PR 29-SEP-2000; 2000US-0236521P.
PR 29-SEP-2000; 2000US-0236522P.
PR 29-SEP-2000; 2000US-0236523P.
PR 29-SEP-2000; 2000US-0236524P.
PR 29-SEP-2000; 2000US-0236525P.
PR 29-SEP-2000; 2000US-0236526P.
PR 29-SEP-2000; 2000US-0236527P.
PR 29-SEP-2000; 2000US-0236528P.
PR 29-SEP-2000; 2000US-0236529P.
PR 29-SEP-2000; 2000US-0236530P.
PR 29-SEP-2000; 2000US-0236531P.
PR 29-SEP-2000; 2000US-0236532P.
PR 29-SEP-2000; 2000US-0236533P.
PR 29-SEP-2000; 2000US-0236534P.
PR 29-SEP-2000; 2000US-0236535P.
PR 29-SEP-2000; 2000US-0236536P.
PR 29-SEP-2000; 2000US-0236537P.
PR 29-SEP-2000; 2000US-0236538P.
PR 29-SEP-2000; 2000US-0236539P.
PR 29-SEP-2000; 2000US-0236540P.
PR 29-SEP-2000; 2000US-0236541P.
PR 29-SEP-2000; 2000US-0236542P.
PR 29-SEP-2000; 2000US-0236543P.
PR 29-SEP-2000; 2000US-0236544P.
PR 29-SEP-2000; 2000US-0236545P.
PR 29-SEP-2000; 2000US-0236546P.
PR 29-SEP-2000; 2000US-0236547P.
PR 29-SEP-2000; 2000US-0236548P.
PR 29-SEP-2000; 2000US-0236549P.
PR 29-SEP-2000; 2000US-0236550P.
PR 29-SEP-2000; 2000US-0236551P.
PR 29-SEP-2000; 2000US-0236552P.
PR 29-SEP-2000; 2000US-0236553P.
PR 29-SEP-2000; 2000US-0236554P.
PR 29-SEP-2000; 2000US-0236555P.
PR 29-SEP-2000; 2000US-0236556P.
PR 29-SEP-2000; 2000US-0236557P.
PR 29-SEP-2000; 2000US-0236558P.
PR 29-SEP-2000; 2000US-0236559P.
PR 29-SEP-2000; 2000US-0236560P.
PR 29-SEP-2000; 2000US-0236561P.
PR 29-SEP-2000; 2000US-0236562P.
PR 29-SEP-2000; 2000US-0236563P.
PR 29-SEP-2000; 2000US-0236564P.
PR 29-SEP-2000; 2000US-0236565P.
PR 29-SEP-2000; 2000US-0236566P.
PR 29-SEP-2000; 2000US-0236567P.
PR 29-SEP-2000; 2000US-0236568P.
PR 29-SEP-2000; 2000US-0236569P.
PR 29-SEP-2000; 2000US-0236570P.
PR 29-SEP-2000; 2000US-0236571P.
PR 29-SEP-2000; 2000US-0236572P.
PR 29-SEP-2000; 2000US-0236573P.
PR 29-SEP-2000; 2000US-0236574P.
PR 29-SEP-2000; 2000US-0236575P.
PR 29-SEP-2000; 2000US-0236576P.
PR 29-SEP-2000; 2000US-0236577P.
PR 29-SEP-2000; 2000US-0236578P.
PR 29-SEP-2000; 2000US-0236579P.
PR 29-SEP-2000; 2000US-0236580P.
PR 29-SEP-2000; 2000US-0236581P.
PR 29-SEP-2000; 2000US-0236582P.
PR 29-SEP-2000; 2000US-0236583P.
PR 29-SEP-2000; 2000US-0236584P.
PR 29-SEP-2000; 2000US-0236585P.
PR 29-SEP-2000; 2000US-0236586P.
PR 29-SEP-2000; 2000US-0236587P.
PR 29-SEP-2000; 2000US-0236588P.
PR 29-SEP-2000; 2000US-0236589P.
PR 29-SEP-2000; 2000US-0236590P.
PR 29-SEP-2000; 2000US-0236591P.
PR 29-SEP-2000; 2000US-0236592P.
PR 29-SEP-2000; 2000US-0236593P.
PR 29-SEP-2000; 2000US-0236594P.
PR 29-SEP-2000; 2000US-0236595P.
PR 29-SEP-2000; 2000US-0236596P.
PR 29-SEP-2000; 2000US-0236597P.
PR 29-SEP-2000; 2000US-0236598P.
PR 29-SEP-2000; 2000US-0236599P.
PR 29-SEP-2000; 2000US-0236600P.
PR 29-SEP-2000; 2000US-0236601P.
PR 29-SEP-2000; 2000US-0236602P.
PR 29-SEP-2000; 2000US-0236603P.
PR 29-SEP-2000; 2000US-0236604P.
PR 29-SEP-2000; 2000US-0236605P.
PR 29-SEP-2000; 2000US-0236606P.
PR 29-SEP-2000; 2000US-0236607P.
PR 29-SEP-2000; 2000US-0236608P.
PR 29-SEP-2000; 2000US-0236609P.
PR 29-SEP-2000; 2000US-0236610P.
PR 29-SEP-2000; 2000US-0236611P.
PR 29-SEP-2000; 2000US-0236612P.
PR 29-SEP-2000; 2000US-0236613P.
PR 29-SEP-2000; 2000US-0236614P.
PR 29-SEP-2000; 2000US-0236615P.
PR 29-SEP-2000; 2000US-0236616P.
PR 29-SEP-2000; 2000US-0236617P.
PR 29-SEP-2000; 2000US-0236618P.
PR 29-SEP-2000; 2000US-0236619P.
PR 29-SEP-2000; 2000US-0236620P.
PR 29-SEP-2000; 2000US-0236621P.
PR 29-SEP-2000; 2000US-0236622P.
PR 29-SEP-2000; 2000US-0236623P.
PR 29-SEP-2000; 2000US-0236624P.
PR 29-SEP-2000; 2000US-0236625P.
PR 29-SEP-2000; 2000US-0236626P.
PR 29-SEP-2000; 2000US-0236627P.
PR 29-SEP-2000; 2000US-0236628P.
PR 29-SEP-2000; 2000US-0236629P.
PR 29-SEP-2000; 2000US-0236630P.
PR 29-SEP-2000; 2000US-0236631P.
PR 29-SEP-2000; 2000US-0236632P.
PR 29-SEP-2000; 2000US-0236633P.
PR 29-SEP-2000; 2000US-0236634P.
PR 29-SEP-2000; 2000US-0236635P.
PR 29-SEP-2000; 2000US-0236636P.
PR 29-SEP-2000; 2000US-0236637P.
PR 29-SEP-2000; 2000US-0236638P.
PR 29-SEP-2000; 2000US-0236639P.
PR 29-SEP-2000; 2000US-0236640P.
PR 29-SEP-2000; 2000US-0236641P.
PR 29-SEP-2000; 2000US-0236642P.
PR 29-SEP-2000; 2000US-0236643P.
PR 29-SEP-2000; 2000US-0236644P.
PR 29-SEP-2000; 2000US-0236645P.
PR 29-SEP-2000; 2000US-0236646P.
PR 29-SEP-2000; 2000US-0236647P.
PR 29-SEP-2000; 2000US-0236648P.
PR 29-SEP-2000; 2000US-0236649P.
PR 29-SEP-2000; 2000US-0236650P.
PR 29-SEP-2000; 2000US-0236651P.
PR 29-SEP-2000; 2000US-0236652P.
PR 29-SEP-2000; 2000US-0236653P.
PR 29-SEP-2000; 2000US-0236654P.
PR 29-SEP-2000; 2000US-0236655P.
PR 29-SEP-2000; 2000US-0236656P.
PR 29-SEP-2000; 2000US-0236657P.
PR 29-SEP-2000; 2000US-0236658P.
PR 29-SEP-2000; 2000US-0236659P.
PR 29-SEP-2000; 2000US-0236660P.
PR 29-SEP-2000; 2000US-0236661P.
PR 29-SEP-2000; 2000US-0236662P.
PR 29-SEP-2000; 2000US-0236663P.
PR 29-SEP-2000; 2000US-0236664P.
PR 29-SEP-2000; 2000US-0236665P.
PR 29-SEP-2000; 2000US-0236666P.
PR 29-SEP-2000; 2000US-0236667P.
PR 29-SEP-2000; 2000US-0236668P.
PR 29-SEP-2000; 2000US-0236669P.
PR 29-SEP-2000; 2000US-0236670P.
PR 29-SEP-2000; 2000US-0236671P.
PR 29-SEP-2000; 2000US-0236672P.
PR 29-SEP-2000; 2000US-0236673P.
PR 29-SEP-2000; 2000US-0236674P.
PR 29-SEP-2000; 2000US-0236675P.
PR 29-SEP-2000; 2000US-0236676P.
PR 29-SEP-2000; 2000US-0236677P.
PR 29-SEP-2000; 2000US-0236678P.
PR 29-SEP-2000; 2000US-0236679P.
PR 29-SEP-2000; 2000US-0236680P.
PR 29-SEP-2000; 2000US-0236681P.
PR 29-SEP-2000; 2000US-0236682P.
PR 29-SEP-2000; 2000US-0236683P.
PR 29-SEP-2000; 2000US-0236684P.
PR 29-SEP-2000; 2000US-0236685P.
PR 29-SEP-2000; 2000US-0236686P.
PR 29-SEP-2000; 2000US-0236687P.
PR 29-SEP-2000; 2000US-0236688P.
PR 29-SEP-2000; 2000US-0236689P.
PR 29-SEP-2000; 2000US-0236690P.
PR 29-SEP-2000; 2000US-0236691P.
PR 2

PR 01-SEP-2000; 2000US-0232987P.
PR 01-SEP-2000; 2000US-02329343P.
PR 01-SEP-2000; 2000US-0229345P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235844P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249220P.
PR 17-NOV-2000; 2000US-0249221P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249248P.
PR 17-NOV-2000; 2000US-0249249P.
PR 17-NOV-2000; 2000US-0249250P.
PR 17-NOV-2000; 2000US-0249251P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-476224/51.
XX
XX
PT Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the respiratory system including respiratory cancers
PT and also for testing and detection e.g. diagnosis.
XX
PS Disclosure; SED ID No 657; 546pp; English.
XX
XX The present invention relates to the isolation of novel human respiratory
CC antigens (AAU17685-AAU17975), and cDNA and genomic sequences encoding for
CC these polypeptides. The sequences of the invention are useful for
CC preventing, treating and/or prognosing disorders related to the
CC respiratory system including throat disorders (e.g. vocal cord paralysis,
CC tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic
CC disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose
CC disorders and cancers of the respiratory tissues e.g. lung cancer. The
CC polynucleotide sequences of the invention are useful in gene therapy and
CC antisense therapy. AA58161-AA528764 represent genomic sequences encoding
CC for novel human respiratory antigens. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 224 BP; 76 A; 39 C; 58 G; 51 T; 0 U; 0 Other;
Query Match 80.0%; Score 19.2; DB 4; Length 224;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CY 1 TCTGGATCTCAGTCATCCACA 24
DB 162 TCTGGCTTCAGTCATCCATCA 139
RESULT 9
ID A0641419 standard; DNA; 224 BP.
XX A0641419;
AC
XX 26-FEB-2004 (first entry)
DT

XX Human respiratory system associated genomic DNA seq id 657.
 XX antiinflammatory; antiallergic; antiasthmatic; cytosol; gene therapy;
 KM respiratory system antigen;
 KM human respiratory system associated polynucleotide;
 KM respiratory system disorder; throat disorder; vocal cord paralysis;
 KM tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
 KM asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
 KM histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
 KM cancer; respiratory tissue cancer; throat cancer; lung cancer;
 KM cancer of the nose; gene therapy; chromosome identification; forensic;
 KM human respiratory system associated protein; ds; human.
 XX Homo sapiens.
 XX OS
 PN US2003215893-A1.
 XX
 PD 20-NOV-2003.
 XX
 PF 07-AUG-2002; 2002US-00212872.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 02-MAR-2000; 2000US-0184664P.
 PR 16-MAR-2000; 2000US-0186350P.
 PR 17-MAR-2000; 2000US-0190874P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209457P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 22-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 25-SEP-2000; 2000US-0234999P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 28-SEP-2000; 2000US-0235935P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244826P.
 PR 01-NOV-2000; 2000US-0244817P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249267P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 01-DEC-2000; 2000US-0250160P.

PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256713P.
PR 06-DEC-2000; 2000US-0251472P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0253678P.
PR 17-JAN-2001; 2001US-00764860.
PR 14-FEB-2002; 2002US-00074095.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-902033/82.
XX
XX Novel respiratory system antigen and polynucleotides encoding the
PT polypeptides, useful for treating diagnosing, treating or preventing
PT consiliitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
PT cancer.
XX
XX Diaclosure; SEQ ID NO 657; 236pp; English.
XX
XX The invention describes an isolated polypeptide (I) comprising an amino
CC acid sequence that is at least 90% identical to polypeptide fragment of
CC any one of 299 respiratory system antigen sequences (PS) and having
CC biological activity, polypeptide domain or epitope of PS, full-length
CC protein of PS, or variant allelic variant or species homolog of PS. (I)
CC or a polynucleotide (II) encoding (I) is also useful for diagnosing a
CC pathological condition or a susceptibility to a pathological condition in
CC a subject which involves determining the presence or absence of mutation
CC in (II) or determining the presence or amount of expression of (I) in a
CC biological sample and diagnosing a pathological condition based on the
CC result. The human respiratory system associated polynucleotides, the
CC polypeptides encoded by them, and antibodies that immunospecifically bind
CC these polypeptides are useful in diagnosis, treatment, prevention and/or
CC prognosis of disorders of respiratory system such as throat disorders
CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
CC (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
CC barcodosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
CC cancer of the nose). The polynucleotides are useful in gene therapy
CC techniques, for chromosome identification, identifying individuals from

Query March 80.0%; Score 19.2; DB 10; Length 224;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCTGGATCTCAGTCATCAACA 24
Db 162 TCTGGCTTTCATGTCATCA 139

RESULT 10
ADG41418/c
ID ADG41418 standard; DNA; 224 BP.
XX
XX ADG41418;
XX
XX 26-FEB-2004 (first entry)
XX
XX Human respiratory system associated genomic DNA seq id 656.
XX
XX antiinflammatory; antiallergic; antiaesthetic; cyrostatic; gene therapy;
KW respiratory system antigen;
KW human respiratory system associated polynucleotide;
KW respiratory system disorder; throat disorder; vocal cord paralysis;

KW consiliitis; laryngitis; lung disorder; pneumonia; allergic disorder;
KW asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
KW histiocytosis; barcodosis; nose disorder; rhinitis; sinusitis; neoplasms;
KW cancer; respiratory tissue cancer; throat cancer; lung cancer;
KW cancer of the nose; gene therapy; chromosome identification; forensic;
KW human respiratory system associated protein; ds; human.
XX
XX Homo sapiens.
XX
XX US2003215893-A1.
PD 20-NOV-2003.
XX
XX 07-AUG-2002; 2002US-00212872.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218230P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 26-JUL-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225269P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0226279P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 21-SEP-2000; 2000US-0234227P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 28-SEP-2000; 2000US-0235935P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0241222P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241825P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246522P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251198P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.

PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251900P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-0076480P.
PR 14-FEB-2002; 2002US-00074095.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI, 2003-902033/82.
DR
XX
XX Novel respiratory system antigen and polynucleotides encoding the
PT polypeptides, useful for treating diagnosing, treating or preventing
PT pneumonia, asthma and cystic fibrosis, emphysema, throat
PT cancer.
XX
XX Disclosure; SEQ ID NO 656; 236bp; English.
XX
XX The invention describes an isolated polypeptide (I) comprising an amino
CC acid sequence that is at least 90% identical to polypeptide fragment of
CC any one of 299 respiratory system antigen sequences (PS) and having
CC biological activity, polypeptide domain or epitope of PS, full-length
CC protein of PS, or variant, allelic variant or species homolog of PS. (I)
CC or a polynucleotide (II) encoding (I) is also useful for diagnosing a
CC pathological condition or a susceptibility to a pathological condition in
CC a subject which involves determining the presence or absence of mutation
CC in (II) or determining the presence or amount of expression of (I) in a
CC biological sample and diagnosing a pathological condition based on the
CC result. The human respiratory system associated polynucleotides, the
CC polypeptides encoded by them, and antibodies that immunospecifically bind
CC these polypeptides are useful in diagnosis, treatment, prevention and/or
CC prognosis of disorders of respiratory system such as throat disorders
CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
CC (e.g., pneumonia), allergic disorders, (e.g., asthma and eosinophilic
CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
CC sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
CC cancer of the nose). The polynucleotides are useful in gene therapy
CC techniques, for chromosome identification, identifying individuals from
Query Match 80.0%; Score 19.2; DB 10; Length 224;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCTGGATCTCAGTCATCCACA 24
Db 162 TCTGGGCTTCAGTCATCCATCA 139
RESULT 11
ADG41417/c
ID ADG41417 standard; DNA; 224 BP.
XX
XX AC ADG41417;
XX
XX 26-FEB-2004 (first entry)
DE Human respiratory system associated genomic DNA seq id 655.
XX
XX antiinflammatory; antiallergic; antiaesthetic; cytostatic; gene therapy;
KM human respiratory system antigen;
KM human respiratory system associated polynucleotide;
KM respiratory system disorder; throat disorder; vocal cord paralysis;
KM tonsillitis; laryngitis; lung disorder; pneumonia; allergic disorder;
KM asthma; eosinophilic pneumonia; pleurisy; cystic fibrosis; emphysema;
KM histiocytosis; sarcoidosis; nose disorder; rhinitis; sinusitis; neoplasm;
KM cancer; respiratory tissue cancer; throat cancer; lung cancer;
KM cancer of the nose; gene therapy; chromosome identification; forensic;
KM human respiratory system associated protein; ds; human.
XX

OS Homo sapiens.
 XX US2003215893-A1.
 XX 20-NOV-2003.
 PF 07-AUG-2002; 2002US-00212872.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198122P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225477P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0228287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 28-SEP-2000; 2000US-0235935P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249255P.
 PR 17-NOV-2000; 2000US-0249257P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251858P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764860.
 PR 14-FEB-2002; 2002US-00074095.

XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-902033/82.
 XX Novel respiratory system antigen and polynucleotides encoding the
 PT polypeptides, useful for treating diagnosing, treating or preventing
 PT tonsillitis, pneumonia, asthma and cystic fibrosis, emphysema, throat
 cancer.
 XX Disclosure; SEQ ID NO 655; 236pp; English.
 XX The invention describes an isolated polypeptide (I) comprising an amino
 CC acid sequence that is at least 90% identical to polypeptide fragment of
 CC any one of 239 respiratory system antigen sequences (PS), full-length
 CC biological activity, polypeptide domain or epitope of PS, full-length
 CC protein of PS, or variant, allelic variant or species homolog of PS. (I)
 CC or a polynucleotide (II) encoding (I) is also useful for diagnosing a
 CC pathological condition or a susceptibility to a pathological condition in
 CC a subject which involves determining the presence or absence of mutation
 CC in (II) or determining the presence or amount of expression of (I) in a
 CC biological sample and diagnosing a pathological condition based on the
 CC result. The human respiratory system associated polynucleotides, the
 CC polypeptides encoded by them, and antibodies that immunospecifically bind
 CC these polypeptides are useful in diagnosis, treatment, prevention and/or
 CC prognosis of disorders of respiratory system such as throat disorders
 CC (e.g., vocal cord paralysis, tonsillitis, and laryngitis), lung disorders
 CC (e.g., pneumonia), allergic disorders (e.g., asthma and eosinophilic
 CC pneumonia), pleurisy, cystic fibrosis, emphysema, histiocytosis,
 CC sarcoidosis, nose disorders (rhinitis and sinusitis), neoplasms and/or
 CC cancers of respiratory tissues (e.g., throat cancer, lung cancer, and
 CC cancer of the nose). The polynucleotides are useful in gene therapy
 CC techniques, for chromosome identification, identifying individuals from

Query Match 80.0%; Score 19.2; DB 10; Length 224;
 Best Local Similarity 87.5%; Pred. No. 27;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TCTGGATCTCAGTCATCCACA 24
 DB 162 TCTGGGCTTTCAGTCATCCACA 139

RESULT 12
 AD197191/c
 ID AD197191 standard; DNA; 224 BP.
 XX AD197191;
 XX 04-NOV-2004 (first entry)
 DE Human respiratory system associated polypeptide-related DNA SeqID55.
 XX
 XX respiratory system-related polypeptide; antiasthmatic; antibacterial;
 KM antiinflammatory; cytotoxic; antianaemic; antiallergic; gene therapy;
 KM pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;
 KM anaemia; leukaemia; inflammation; sinusitis;
 KM chronic obstructive pulmonary disease; infectious disease; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN US200307704-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 14-FEB-2002; 2002US-00074095.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 14-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0237032P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.

PR	13-OCT-2000	2000US-0239935P.
PR	13-OCT-2000	2000US-0239937P.
PR	20-OCT-2000	2000US-0240960P.
PR	20-OCT-2000	2000US-024121P.
PR	20-OCT-2000	2000US-0241785P.
PR	20-OCT-2000	2000US-0241786P.
PR	20-OCT-2000	2000US-0241787P.
PR	20-OCT-2000	2000US-0241808P.
PR	20-OCT-2000	2000US-0241809P.
PR	20-OCT-2000	2000US-0241826P.
PR	01-NOV-2000	2000US-024617P.
PR	08-NOV-2000	2000US-0246474P.
PR	08-NOV-2000	2000US-0246475P.
PR	08-NOV-2000	2000US-0246476P.
PR	08-NOV-2000	2000US-0246522P.
PR	08-NOV-2000	2000US-0246528P.
PR	08-NOV-2000	2000US-0246532P.
PR	08-NOV-2000	2000US-0246509P.
PR	08-NOV-2000	2000US-0246610P.
PR	08-NOV-2000	2000US-0246611P.
PR	08-NOV-2000	2000US-0246912P.
PR	17-NOV-2000	2000US-0249213P.
PR	17-NOV-2000	2000US-0249214P.
PR	17-NOV-2000	2000US-0249208P.
PR	17-NOV-2000	2000US-0249209P.
PR	17-NOV-2000	2000US-0249210P.
PR	17-NOV-2000	2000US-0249211P.
PR	17-NOV-2000	2000US-0249217P.
PR	17-NOV-2000	2000US-0249218P.
PR	17-NOV-2000	2000US-0249244P.
PR	17-NOV-2000	2000US-0249245P.
PR	17-NOV-2000	2000US-0249264P.
PR	17-NOV-2000	2000US-0249265P.
PR	17-NOV-2000	2000US-0249297P.
PR	17-NOV-2000	2000US-0249299P.
PR	17-NOV-2000	2000US-0249300P.
PR	01-DEC-2000	2000US-0250160P.
PR	01-DEC-2000	2000US-0250391P.
PR	05-DEC-2000	2000US-0251030P.
PR	05-DEC-2000	2000US-0251988P.
PR	05-DEC-2000	2000US-0256719P.
PR	06-DEC-2000	2000US-025179P.
PR	08-DEC-2000	2000US-0251856P.
PR	08-DEC-2000	2000US-0251868P.
PR	08-DEC-2000	2000US-0251899P.
PR	08-DEC-2000	2000US-0251990P.
PR	11-DEC-2000	2000US-0254097P.
PR	05-JAN-2001	2001US-0259678P.
PR	17-JAN-2001	2001US-00764860.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Rosen CA, Ruben SM, Barash SC,	
XX		
DR	WPI, 2003-765403/72.	
XX		
PT	New human respiratory system-related polypeptide and genes, useful for	
XX	treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic	
PT	fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or	
XX	sinusitis.	
XX		
XX	Disclosure; SEQ ID NO 655; 202pp; English.	

CC	This invention is related to a novel isolated polypeptide, which comprises a human respiratory system-related polypeptide, and the DNA sequence which encodes it. The invention may be useful for the development of compounds with an antiasthmatic, antibacterial,
CC	antinfl ammatory, cytosstatic, anti anaemic or antiallergic activity. In addition, the sequences disclosed may be useful for gene therapy. The
CC	polypeptide or polynucleotide is useful for treating, preventing or ameliorating a medical condition, for example pneumonia, lung cancer,
CC	cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,
CC	inflammations, sinusitis, chronic obstructive pulmonary disease or infectious diseases. The polypeptide or polynucleotide is also useful for diagnosing any of these diseases or a susceptibility to the disease. The present sequence is that of a human DNA sequence which is related to a human respiratory system associated gene of the invention.
XX	
SQ	. Sequence 224 BP; 76 A; 39 C; 58 G; 51 T; 0 U; 0 Other;
Query Match	80.0%; Score 19.2; DB 11; Length 224;
Best Local Similarity	87.5%; Pred. No. 27;
Matches	21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY	1 TTGGGATCTAGTGCATCCACA 24 DB 162 TGTGGCGTTTCAGTGCAATCATCA 139
RESULT 13	
ID ADI97192/C	
ID ADI97192 standard; DNA; 224 BP.	
AC AC	
AD I97192;	
DT DT	
XX XX	04-NOV-2004 (first entry)
DE DE	Human respiratory system associated polypeptide-related DNA SeqID56.
KX KX	respiratory system-related polypeptide; antiaesthetic; antibacterial; KW KW antimflammatory; cytostatic; antifla emic; antiallergic; gene therapy; KM KM pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis; KK KK anemia; leukaemia; inflammation; sinusitis; XX XX chronic obstructive pulmonary disease; infectious disease; human; ds. OS OS Homo sapiens. CS CS PN PN US200307704-A1. PD PD 24-APR-2003. PF PF 14-FEB-2002; 2002US-00074095. XX XX PR PR 31-JAN-2000; 2000US-0179065P. PR PR 04-FEB-2000; 2000US-0180628P. PR PR 24-FEB-2000; 2000US-0184664P. PR PR 02-MAR-2000; 2000US-0186350P. PR PR 16-MAR-2000; 2000US-0189874P. PR PR 17-MAR-2000; 2000US-0190076P. PR PR 18-APR-2000; 2000US-0198123P. PR PR 19-MAY-2000; 2000US-0205515P. PR PR 07-JUN-2000; 2000US-0209467P. PR PR 28-JUN-2000; 2000US-0214886P. PR PR 30-JUN-2000; 2000US-0215135P. PR PR 07-JUL-2000; 2000US-0216647P. PR PR 07-JUL-2000; 2000US-0216880P. PR PR 11-JUL-2000; 2000US-0217497P. PR PR 11-JUL-2000; 2000US-0217496P. PR PR 14-JUL-2000; 2000US-0218290P. PR PR 26-JUL-2000; 2000US-0230963P. PR PR 26-JUL-2000; 2000US-0230964P. PR PR 14-AUG-2000; 2000US-0224518P. PR PR 14-AUG-2000; 2000US-0224519P. PR PR 14-AUG-2000; 2000US-0225213P. PR PR 14-AUG-2000; 2000US-0225214P.

PR	14-AUG-2000	2000US-0225266P	
PR	14-AUG-2000	2000US-0225267P	
PR	14-AUG-2000	2000US-0225268P	
PR	14-AUG-2000	2000US-0225270P	
PR	14-AUG-2000	2000US-0225447P	
PR	14-AUG-2000	2000US-0225757P	
PR	14-AUG-2000	2000US-0225758P	
PR	14-AUG-2000	2000US-0225759P	
PR	16-AUG-2000	2000US-0226292P	
PR	22-AUG-2000	2000US-0226619P	
PR	22-AUG-2000	2000US-0226688P	
PR	22-AUG-2000	2000US-0227182P	
PR	30-AUG-2000	2000US-0227009P	
PR	30-AUG-2000	2000US-02288924P	
PR	01-SEP-2000	2000US-0229287P	
PR	01-SEP-2000	2000US-0229333P	
PR	01-SEP-2000	2000US-0229344P	
PR	01-SEP-2000	2000US-0229345P	
PR	05-SEP-2000	2000US-0229509P	
PR	05-SEP-2000	2000US-0229533P	
PR	06-SEP-2000	2000US-0230437P	
PR	06-SEP-2000	2000US-0230488P	
PR	08-SEP-2000	2000US-0231242P	
PR	08-SEP-2000	2000US-0231243P	
PR	08-SEP-2000	2000US-0231244P	
PR	08-SEP-2000	2000US-0231413P	
PR	08-SEP-2000	2000US-0231414P	
PR	08-SEP-2000	2000US-0232080P	
PR	08-SEP-2000	2000US-0232081P	
PR	12-SEP-2000	2000US-0231968P	
PR	14-SEP-2000	2000US-0232397P	
PR	14-SEP-2000	2000US-0232398P	
PR	14-SEP-2000	2000US-0232399P	
PR	14-SEP-2000	2000US-0232401P	
PR	14-SEP-2000	2000US-0232402P	
PR	14-SEP-2000	2000US-0233063P	
PR	14-SEP-2000	2000US-0233064P	
PR	14-SEP-2000	2000US-0233065P	
PR	21-SEP-2000	2000US-0234223P	
PR	21-SEP-2000	2000US-0234224P	
PR	25-SEP-2000	2000US-0234987P	
PR	25-SEP-2000	2000US-0235469P	
PR	26-SEP-2000	2000US-0235484P	
PR	27-SEP-2000	2000US-0235834P	
PR	27-SEP-2000	2000US-0235836P	
PR	29-SEP-2000	2000US-0236337P	
PR	29-SEP-2000	2000US-0236367P	
PR	29-SEP-2000	2000US-0236568P	
PR	29-SEP-2000	2000US-0236569P	
PR	29-SEP-2000	2000US-0236730P	
PR	02-OCT-2000	2000US-02376802P	
PR	02-OCT-2000	2000US-02370387P	
PR	02-OCT-2000	2000US-02370388P	
PR	02-OCT-2000	2000US-0237039P	
PR	02-OCT-2000	2000US-0237040P	
PR	13-OCT-2000	2000US-0239395P	
PR	13-OCT-2000	2000US-0239937P	
PR	20-OCT-2000	2000US-0240960P	
PR	20-OCT-2000	2000US-0241211P	
PR	20-OCT-2000	2000US-0241785P	
PR	20-OCT-2000	2000US-0241786P	
PR	20-OCT-2000	2000US-0241787P	
PR	20-OCT-2000	2000US-0241808P	
PR	20-OCT-2000	2000US-0241809P	
PR	20-OCT-2000	2000US-0241866P	
PR	01-NOV-2000	2000US-0244617P	
PR	08-NOV-2000	2000US-0246474P	
PR	08-NOV-2000	2000US-0246475P	
PR	08-NOV-2000	2000US-0246477P	
PR	08-NOV-2000	2000US-0246478P	
PR	08-NOV-2000	2000US-0246533P	
PR	08-NOV-2000	2000US-0246534P	

PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246539P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251858P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254077P.
PR	05-JAN-2001;	2001US-0259676P.
PR	17-JAN-2001;	2001US-00764860.
PA	(HUMA-)	HUMAN GENOME SCI INC.
PL	Rosen CA,	Ruben SM, Barash SC;
XX	DR	WPI, 2003-765403/72.
PT	PT	New human respiratory system-related polypeptide and genes, useful for
PT	PT	treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic
PT	PT	fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or
PT	PT	sinusitis.
PS	XX	Disclosure; SEQ ID NO 656; 202pp; English.
CC	CC	This invention is related to a novel isolated polypeptide, which
CC	CC	comprises a human respiratory system-related polypeptide, and the DNA
CC	CC	sequence which encodes it. The invention may be useful for the
CC	CC	development of compounds with an antiasthmatic, antibacterial,
CC	CC	antiinflammatory, cytostatic, antitumoric or antiallergic activity. In
CC	CC	addition, the sequences disclosed may be useful for gene therapy. The
CC	CC	polypeptide or polynucleotide is useful for treating, preventing or
CC	CC	ameliorating a medical condition, for example pneumonia, lung cancer,
CC	CC	cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,
CC	CC	inflammations, sinusitis, chronic obstructive pulmonary disease or
CC	CC	infectious diseases. The polypeptide or polynucleotide is also useful for
CC	CC	diagnosing any of these diseases or a susceptibility to the disease. The
CC	CC	present sequence is that of a human DNA sequence which is related to a
CC	CC	human respiratory system associated gene of the invention.
XX	XX	Sequence 224 BP; 76 A; 39 C; 58 G;51 T; 0 U; 0 Other;

Query Match 80.0%; Score 19.2; DB 11; Length 224;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCTGGATTCAGTCATCCACA 24
162 TCTGGGCTTCAGTCATCCATCA 139
Db
RESULT 14
AD197193/c
ID AD197193 standard; DNA; 224 BP.
XX
AC AD197193;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human respiratory system associated polypeptide-related DNA SegID657.
XX
KW respiratory system-related polypeptide; antiasthmatic; antibacterial;
KW antiinflammatory; cytotoxic; antianaemic; antiallergic; gene therapy;
KW pneumonia; lung cancer; cystic fibrosis; asthma; sarcoidosis; rhinitis;
KW anaemia; leukaemia; inflammation; sinusitis;
KW chronic obstructive pulmonary disease; infectious disease; human; ds.
XX
OS Homo sapiens.
XX
PN US2003077704-A1.
XX
PD 24-APR-2003.
XX
PF 14-FEB-2002; 2002US-00074095.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220964P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231124P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236380P.
PR 29-SEP-2000; 2000US-0236381P.
PR 29-SEP-2000; 2000US-0236382P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241121P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246530P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.

XX	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249219P.
PR	17-NOV-2000;	2000US-0249220P.
PR	17-NOV-2000;	2000US-0249221P.
PR	17-NOV-2000;	2000US-0249222P.
PR	17-NOV-2000;	2000US-0249223P.
PR	17-NOV-2000;	2000US-0249224P.
PR	17-NOV-2000;	2000US-0249225P.
PR	17-NOV-2000;	2000US-0249226P.
PR	17-NOV-2000;	2000US-0249227P.
PR	17-NOV-2000;	2000US-0249228P.
PR	17-NOV-2000;	2000US-0249229P.
PR	17-NOV-2000;	2000US-0249230P.
PR	17-NOV-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0254739P.
PR	08-DEC-2000;	2000US-0251855P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0254979P.
PR	11-DEC-2000;	2000US-0256797P.
PR	05-JAN-2001;	2001US-0256797P.
PR	17-JAN-2001;	2001US-00764860.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
P1	Rosen CA,	Ruben SM, Barash SC;
XX		
DR	WPI;	2003-765403/72.
XX		
PT	New human respiratory system-related polypeptide and genes, useful for	
PT	treating, preventing or diagnosing e.g. pneumonia, lung cancer, cystic	
PT	fibrosis, asthma, sarcoidosis, rhinitis, leukemia, inflammations or	
PT	sinusitis.	
XX		
PS	Disclosure:	SEQ ID NO 657; 202bp; English.
XX		
CC	This invention is related to a novel isolated polypeptide, which	
CC	comprises a human respiratory system-related polypeptide, and the DNA	
CC	sequence which encodes it. The invention may be useful for the	
CC	development of compounds with an antiallergic, antibacterial,	
CC	antiinflammatory, cytostatic, antineoplastic or antiallergic activity. In	
CC	addition, the sequences disclosed may be useful for gene therapy. The	
CC	polypeptide or polynucleotide is useful for treating, preventing or	
CC	ameliorating a medical condition, for example pneumonia, lung cancer,	
CC	cystic fibrosis, asthma, sarcoidosis, rhinitis, anaemia, leukaemia,	
CC	inflammatory, sinusitis, chronic obstructive pulmonary disease or	
CC	infectious diseases. The polypeptide or polynucleotide is also useful for	
CC	diagnosing any of these diseases or a susceptibility to the disease. The	
CC	present sequence is that of a human DNA sequence which is related to a	
CC	human respiratory system associated gene of the invention.	
XX		
XX	Sequence	224 BP; 76 A; 39 C; 58 G; 51 T; 0 U; 0 Other;
XX		
Query Match		80.0%; Score 19.2; DB 11; Length 224;
Best Local Similarity		87.5%; Pred. No. 27;
Matches	21; Conservative	0; Mismatches 3; Indels 0; Gaps 0
QY	1 TCTGGGATCTCAGTCATCCACA	24
DB	162 TCTGGGCTTTCAGTCATCCATCA	139
RESULT 15		
AAQ48675/c		
ID	AAQ48675 standard; cDNA, 3323 BP.	
XX		
AC	AAQ48675;	
XX		
XX	25-MAR-2003 (revised)	
DT	25-FEB-1994 (first entry)	
XX		

```

DE Human myotonic dystrophy (DM) gene.
XX Abnormality; muscular dystrophy; CHR 19; chromosome 19; protein kinase;
XX adult; ss.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 1..1749
FT /*tag= a
FT /note= "coding region"
FT repeat_region 1972..2004
FT /*tag= b
FT /note= "CTG repeat region"
XX
XX
XX W09317104-A1.
XX
XX 02-SEP-1993.
XX
XX 19-FEB-1993; 93WO-US001545.
XX
XX 20-FEB-1992; 92US-00839255.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Brook JD, Housman DE;
XX
XX WPI; 1993-288410/36.
XX
XX P-PSDB; AAR41000.
XX
XX DNA sequence of myotonic dystrophy gene - used to produce probes and
XX identify CHR 19 abnormality and protein kinase reponsible.
XX
XX Claim 3; Fig 5; 64bp; English.
XX
XX
XX The sequence is that of the adult human brain cDNA clone C28 which
XX comprises the "myotonic dystrophy gene" which includes in its 3'
XX untranslated region a transcribed Ctg triplet repeat. The sequence
XX encodes a protein kinase, detection of expression of this protein kinase
XX may be used as a method for determining whether an individual is likely
XX to be affected with myotonic dystrophy. (Updated on 25-MAR-2003 to
XX correct PW field.)
XX
XX
XX Sequence 3323 BP; 590 A; 1021 C; 1062 G; 608 T; 0 U; 42 Other;
XX
XX
XX Query Match 80.0%; Score 19.2; DB 2; Length 3323;
XX Best Local Similarity 87.5%; Pred. No. 44;
XX Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 TCTGGGATCTCAGTCATCCACCA 24
XX |||||
XX 322 TCGGGGCTCAGTCATCCAAA 3199

```


1	10966	100.0	10966	6	E48993	Creation of
2	10966	100.0	10966	6	BD012010	ABD012010
3	10966	100.0	10966	8	AB024006	Hordeum v
4	971.8	8.9	1895	8	AB005788	AB005788 Hordeum v
5	744.6	6.8	1660	8	D88273	D88273 Hordeum vul
6	366.8	3.3	126315	2	AP004116	AP004116 Oryza sat
7	366.8	3.3	159075	2	AP005743	AP005743 Oryza sat
8	326	3.0	1035	8	BT009504	BT009504 Triticum
9	273.8	2.5	14216	8	AP005310	AP005310 Oryza sat
10	273.8	2.5	184760	8	AP005532	AP005532 Oryza sat
11	213	1.9	111173	8	AP004679	AP004679 Oryza sat
12	213	1.9	147640	2	AP003542	AP003542 Oryza sat
13	212.6	1.9	93695	8	ATP2009	ATL021149 Arabidops
14	212.6	1.9	199749	8	ATP2009	ATL015172 Arabidops
15	209.6	1.9	147540	8	AC150658	AC150658 Oryza sat
16	209.6	1.9	148785	8	AC134624	AC134624 Oryza sat
17	194.4	1.8	61384	8	AC007048	AC007048 Arabidops
18	192	1.8	73840	8	AB007644	AB007644 Arabidops
19	184.8	1.7	610	11	AL773360	AL773360 Arabidops

AC006585	Arabidops
AC005394	Arabidops
AL161559	Arabidops
AB018112	Arabidops
AK654009	Sequence
AC060537	Oryza sat
AY142527	Arabidops
AL166113	Arabidops
AK412376	Sequence
AY507173	Sequence
AK651743	Sequence
BT006593	Arabidops
BT002475	Arabidops
AY085124	Arabidops
AK653614	Sequence
AK108575	Oryza sat
AK575729	Sequence
AK107186	Oryza sat
AK653127	Sequence
AY057289	Sequence
AY091293	Arabidops
AY054204	Arabidops
AF301198	Arabidops
AY050987	Arabidops
AY450993	Colus bl
AK575731	Sequence

ALIGNMENTS

RESULT 1

LOCUS	EA9893	10966 bp	DNA	linear	PAT 31-JAN-2002
DEFINITION	Creation of iron-deficiency resistant rice plant.				
ACCESSION	EA9893				
VERSION	EA9893.1	GI:18629372			
KEYWORDS	JP 2001017012-A/1.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 10966)				
AUTHORS	Mori,S., Nakamishi,N., Takahashi,M. and Nishizawa,N.				
TITLE	Creation of iron-deficiency resistant rice plant				
JOURNAL	Patent: JP 2001017012-A 1 23-JAN-2001				

COMMENT

PN	JP 2001017012-A/1
PD	23-JAN-2001
PF	05-JUL-1999 JP 1999190318

PC	Key	Location/Qualifiers	Source
PC A01N5/00, C12N5/10, C12N5/09//C12N9/10, C12N5/00, C12N15/00	Key	Location/Qualifiers	Source
FT	1.10966		
FT	/organism='Hornuudum vulgare L. var. Igr'		

FEATURES

```
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
```

ORIGIN

Query March Similarity	100.0%	Score 10966	DB 6	length 10966
Best Local Similarity	100.0%	Pred. No. 0		
Matches 10966; Conservative	0	Mismatches	0	Indels 0; Gaps 0

Qy 1 CTGATCCCATTCGCAATGATGATTACCTATCAACGAAAGAAAGATGGCATGTGCC 60
Db 1 CTGATCCCATTCGCAATGATGATTACCTATCAACGAAAGAAAGATGGCATGTGCC 60
Qy 61 CTGTGTGTCATCCCTCACTGCGCTTGGCGAATGGCGATACGAGTTAGGTAGAGTTT 120

Db 61 CTGTGTGTCATCCCTCACTGGCTTGCGAATGCGCATCCGAGTTAGGTAAGTGTCTT 120
Qy 121 TTACATGATGTCGTCCGGGCACTGCGAAGAAATCGCTGTGAGGGGACTGAGAGAGTT 180
Db 121 TTACATGATGTCGTCCGGGCACTGCGAAGAAATCGCTGTGAGGGGACTGAGAGAGTT 180
Qy 181 GAGCGATGATGTCGTCGTGTGATGAGCGGAGCTGATGAGTGTCTCACTAACTGAACCAATCA 240
Db 181 GAGCGATGATGTCGTCGTGTGATGAGCGGAGCTGATGAGTGTCTCACTAACTGAACCAATCA 240
Qy 241 GCATTTGGTGTGATGAGTCGAGGAAGCATATGCTTTCGTGCTCCCGATCCGCTTATCTTT 300
Db 241 GCATTTGGTGTGATGAGTCGAGGAAGCATATGCTTTCGTGCTCCCGATCCGCTTATCTTT 300
Qy 301 TTCTCCCAATTTTAAAGAGGATAGATGATGATGCTGCTGGTTGGGTAAGTACGTC 360
Db 301 TTCTCCCAATTTTAAAGAGGATAGATGATGATGCTGCTGGTTGGGTAAGTACGTC 360
Qy 361 ATAGAACCAAGCGAGGCGCGAAATATGCTGGGAGATATGCTGGCAGGCGCGCAACGCG 420
Db 361 ATAGAACCAAGCGAGGCGCGAAATATGCTGGGAGATATGCTGGCAGGCGCGCAACGCG 420
Qy 421 CAGCCCGTCACTGCTGCGAGCGGCGTGCAGAGCGTGCAGAGCGTGCAGCGCGCTGCT 480
Db 421 CAGCCCGTCACTGCTGCGAGCGGCGTGCAGAGCGTGCAGAGCGTGCAGCGCGCTGCT 480
Qy 481 TCTTGTCTGCGGCGCCGGTTCGTGTGCGGTCAAGACACGCTATATAGAACCGTCAATC 540
Db 481 TCTTGTCTGCGGCGCCGGTTCGTGTGCGGTCAAGACACGCTATATAGAACCGTCAATC 540
Qy 541 ACCGCTACTCAATCCGTCCCACTGTTTCTATATACCGCTATAGTAATTCCTGCT 600
Db 541 ACCGCTACTCAATCCGTCCCACTGTTTCTATATACCGCTATAGTAATTCCTGCT 600
Qy 601 GTAGTCTAGTATGTCCT 660
Db 601 GTAGTCTAGTATGTCCT 660
Qy 661 ACCGCAAGCGAGCGAGTTCGCGCGAAGCGGCTTGCGGTGCGCGACCGCGAAGCGCAA 720
Db 661 ACCGCAAGCGAGCGAGTTCGCGCGAAGCGGCTTGCGGTGCGCGACCGCGAAGCGCAA 720
Qy 721 GAGCAACGCGCATGCGTGTGCTGCGCGTGTGAAAGCGCAAGCAACGCGCATGCGTGA 780
Db 721 GAGCAACGCGCATGCGTGTGCTGCGCGTGTGAAAGCGCAAGCAACGCGCATGCGTGA 780
Qy 781 TGCCGATGCGCAACGCGCAAGCAACGCGCATGCGTGTGCGCGTGTGCGCAACGCGCAAG 840
Db 781 TGCCGATGCGCAACGCGCAAGCAACGCGCATGCGTGTGCGCGTGTGCGCAACGCGCAAG 840
Qy 841 CAACGCGCATGCGCAAGCAACGCGCATGCGTGTGCGCGTGTGCGCAACGCGCAAGCAAG 900
Db 841 CAACGCGCATGCGCAAGCAACGCGCATGCGTGTGCGCGTGTGCGCAACGCGCAAGCAAG 900
Qy 901 CAACGCGCATGCGCAAGCAACGCGCATGCGTGTGCGCGTGTGCGCAACGCGCAAGCAAG 960
Db 901 CAACGCGCATGCGCAAGCAACGCGCATGCGTGTGCGCGTGTGCGCAACGCGCAAGCAAG 960
Qy 961 CGAGCATGCGCAAGCAACGCGCATGCGTGTGCGCGTGTGCGCAACGCGCAAGCAAGCAAG 1020
Db 961 CGAGCATGCGCAAGCAACGCGCATGCGTGTGCGCGTGTGCGCAACGCGCAAGCAAGCAAG 1020
Qy 1021 AGAGGATGAGGAGCGGTCGATGGAATTTTCGCGGTGCGCAAGCAAGCAAGCAAGCAAG 1080
Db 1021 AGAGGATGAGGAGCGGTCGATGGAATTTTCGCGGTGCGCAAGCAAGCAAGCAAGCAAG 1080
Qy 1081 GACGCGGCGCAACATGAGCATCCGCGGATGCGGATCAAGATCAGCGCGAGCGTGAAGA 1140
Db 1081 GACGCGGCGCAACATGAGCATCCGCGGATGCGGATCAAGATCAGCGCGAGCGTGAAGA 1140
Qy 1141 GAAGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 GAAGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

Db 1141 GAAGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy 1201 CCGGACGCGCGTGAAGCGCGAGACGCGCGTGCAGCGCGCGCTGCGCAACGCGCGCTTCA 1260
Db 1201 CCGGACGCGCGTGAAGCGCGAGACGCGCGTGCAGCGCGCGCTGCGCAACGCGCGCTTCA 1260
Qy 1261 CTGCTACCCCGCGCGCTGCGCGTCCCGCGCGAGCAAGAGTAAACACACACACACAA 1320
Db 1261 CTGCTACCCCGCGCGCTGCGCGTCCCGCGCGAGCAAGAGTAAACACACACACACAA 1320
Qy 1321 GAACAATTTCTTTTGGCGGT 1380
Db 1321 GAACAATTTCTTTTGGCGGT 1380
Qy 1381 TTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Db 1381 TTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Qy 1441 GACCTTCTCCACCTTATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 1500
Db 1441 GACCTTCTCCACCTTATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 1500
Qy 1501 GTGCGAGGCGGTGCGCGTCAATGCTATGCGCGAGAGCTTCTCTCAACGCGCGCGGAC 1560
Db 1501 GTGCGAGGCGGTGCGCGTCAATGCTATGCGCGAGAGCTTCTCTCAACGCGCGCGGAC 1560
Qy 1561 CAGGCGATGAGATCAATATCCCGGTGCTGCGCGAGAGCTTCTCTCAACGCGCGCGGAC 1620
Db 1561 CAGGCGATGAGATCAATATCCCGGTGCTGCGCGAGAGCTTCTCTCAACGCGCGCGGAC 1620
Qy 1621 CCGGAGCGAGGCTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 1680
Db 1621 CCGGAGCGAGGCTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 1680
Qy 1681 GCAATTTGCACTGATCCCGGCAAGGCGTGGAGATGCAATGCACTGCTGGAATTCAT 1740
Db 1681 GCAATTTGCACTGATCCCGGCAAGGCGTGGAGATGCAATGCACTGCTGGAATTCAT 1740
Qy 1741 CGCGCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 1800
Db 1741 CGCGCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 1800
Qy 1801 TTACTCTACGACATCTGTCCAAAGTTTCAATCTTGTGCTGTAATGATGATCA 1860
Db 1801 TTACTCTACGACATCTGTCCAAAGTTTCAATCTTGTGCTGTAATGATGATGATCA 1860
Qy 1861 GTTCACTGACCTGTGTAATCTTTTGTCCAAATGCAATGCACTGATGTTGCTCAATTA 1920
Db 1861 GTTCACTGACCTGTGTAATCTTTTGTCCAAATGCAATGCACTGATGTTGCTCAATTA 1920
Qy 1921 GGTGCGGAGGTGCGCAAGGCTGCGAATAATGCTGAATGCTGCAAGGATATACGCA 1980
Db 1921 GGTGCGGAGGTGCGCAAGGCTGCGAATAATGCTGAATGCTGCAAGGATATACGCA 1980
Qy 1981 GCTGTCTGCGGAGGCGCGCGCTGCTCAATCCCAATGGAAGTGTGGCAATCAACCTGT 2040
Db 1981 GCTGTCTGCGGAGGCGCGCGCTGCTCAATCCCAATGGAAGTGTGGCAATCAACCTGT 2040
Qy 2041 GCTGTCTGCGGAGGCTGCTGCAAGTCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2100
Db 2041 GCTGTCTGCGGAGGCTGCTGCAAGTCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2100
Qy 2101 AGCGGTGTAAGACCCGAGAAAGATCTTAAAGGAACTTAAAGTAACTTAACTTAACTA 2160
Db 2101 AGCGGTGTAAGACCCGAGAAAGATCTTAAAGGAACTTAAAGTAACTTAACTTAACTA 2160
Qy 2161 TTCTTTTCAATGCTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 2220
Db 2161 TTCTTTTCAATGCTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 2220
Qy 2221 TTGTAAGCAATTTTGT 2280
Db 2221 TTGTAAGCAATTTTGT 2280

QY 2281 CAGACCAGCACTTCACTGAGTCACTCTTGGTATTTACCTGCTTCAAGAAATPAA 2340
 Db 2281 CAGACCAGCACTTCACTGAGTCACTCTTGGTATTTACCTGCTTCAAGAAATPAA 2340
 QY 2341 GTCTTTGGTATTTACCTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2400
 Db 2341 GTCTTTGGTATTTACCTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2400
 QY 2401 CAGGTGATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2460
 Db 2401 CAGGTGATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2460
 QY 2461 ACATTTGATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2520
 Db 2461 ACATTTGATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2520
 QY 2521 CTGAACCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2580
 Db 2521 CTGAACCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2580
 QY 2581 CATGTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2640
 Db 2581 CATGTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2640
 QY 2641 ACATTTGATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2700
 Db 2641 ACATTTGATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2700
 QY 2701 ATTTGTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2760
 Db 2701 ATTTGTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2760
 QY 2761 GCAGCTCTTCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2820
 Db 2761 GCAGCTCTTCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2820
 QY 2821 CTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2880
 Db 2821 CTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2880
 QY 2881 TGTCTCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2940
 Db 2881 TGTCTCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 2940
 QY 2941 AAATCTTGAAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3000
 Db 2941 AAATCTTGAAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3000
 QY 3001 ATGTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3060
 Db 3001 ATGTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3060
 QY 3061 TGAAGTCACTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3120
 Db 3061 TGAAGTCACTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3120
 QY 3121 CTGTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3180
 Db 3121 CTGTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3180
 QY 3181 ACACAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3240
 Db 3181 ACACAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3240
 QY 3241 GTATTAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3300
 Db 3241 GTATTAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3300
 QY 3301 TGAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3360
 Db 3301 TGAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3360

QY 3361 GAATCAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3420
 Db 3361 GAATCAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3420
 QY 3421 CATTAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3480
 Db 3421 CATTAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3480
 QY 3481 AAGCTTGAAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3540
 Db 3481 AAGCTTGAAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3540
 QY 3541 TTGACGTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3600
 Db 3541 TTGACGTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3600
 QY 3601 TCAATTAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3660
 Db 3601 TCAATTAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3660
 QY 3661 TTTGTCAGTCACTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3720
 Db 3661 TTTGTCAGTCACTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3720
 QY 3721 CCTCTCAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3780
 Db 3721 CCTCTCAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3780
 QY 3781 TGTTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3840
 Db 3781 TGTTCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3840
 QY 3841 GAAGAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3900
 Db 3841 GAAGAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3900
 QY 3901 CAGTATCCCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3960
 Db 3901 CAGTATCCCATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 3960
 QY 3961 CAACATCTCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4020
 Db 3961 CAACATCTCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4020
 QY 4021 GTTATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4080
 Db 4021 GTTATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4080
 QY 4081 GGGGGGGGGGCTTGAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4140
 Db 4081 GGGGGGGGGGCTTGAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4140
 QY 4141 CTTTGAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4200
 Db 4141 CTTTGAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4200
 QY 4201 AAGCAAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4260
 Db 4201 AAGCAAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4260
 QY 4261 GGATCAGTGAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4320
 Db 4261 GGATCAGTGAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4320
 QY 4321 CTCAATCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4380
 Db 4321 CTCAATCTGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4380
 QY 4381 CTCTGTCAAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4440
 Db 4381 CTCTGTCAAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4440
 QY 4441 AAGAGAAATGTCAGAGTCACTGCTTGGTATTTGGTCCGATCCCTATGTGTGAGCAACC 4500

```

Db 4441 AAGAGAGAACCCCTCACAAACCTTGTGAGCCCCCAATCATGTTAGCCCGCG 4500
Qy 4501 TCGAGAGCCGGAGTGTCAACATGCGAGCAACCACTCCACCGGGGAGCACTGAC 4560
Db 4501 TCGAGAGCCGGAGTGTCAACATGCGAGCAACCACTCCACCGGGGAGCACTGAC 4560
Qy 4561 CGATTACTGAATATCCCAACCAATCTTAATTTGTGAACAAAATTTAAACAGAAACA 4620
Db 4561 CGATTACTGAATATCCCAACCAATCTTAATTTGTGAACAAAATTTAAACAGAAACA 4620
Qy 4621 ATTTTGAATTTGTGAACAAATTTTAAACGGGTATCTGAACTTTTCAAATTTGT 4680
Db 4621 ATTTTGAATTTGTGAACAAATTTTAAACGGGTATCTGAACTTTTCAAATTTGT 4680
Qy 4681 GATCAAAATTTTAAACGACTCTTCTCAATTTGAGCATATTTAAATTAATAAAA 4740
Db 4681 GATCAAAATTTTAAACGACTCTTCTCAATTTGAGCATATTTAAATTAATAAAA 4740
Qy 4741 GTTCAACAAATTTTGAACCTTTTAAATAATAGCGAACATTTGAAATTTCTAATTTT 4800
Db 4741 GTTCAACAAATTTTGAACCTTTTAAATAATAGCGAACATTTGAAATTTCTAATTTT 4800
Qy 4801 CGAATTTGGAACATTTTCTAATTTCTGAACAAAATTTGAAATTAAGAAATTTTGA 4860
Db 4801 CGAATTTGGAACATTTTCTAATTTCTGAACAAAATTTGAAATTAAGAAATTTTGA 4860
Qy 4861 ATAAATTTTGGAAATGCGATTTTGTGAATTTTGTGAACATATTTTGAATAAA 4920
Db 4861 ATAAATTTTGGAAATGCGATTTTGTGAATTTTGTGAACATATTTTGAATAAA 4920
Qy 4921 CTTTAAAGGTAAATATAATAATAATAATAATAATAATAATAATAATAATAATA 4980
Db 4921 CTTTAAAGGTAAATATAATAATAATAATAATAATAATAATAATAATAATAATA 4980
Qy 4981 AAAGAAATCCGAGAAAGCCCACTGGGAATAGCAATGGAATAAACCAGCCGTCCGCGC 5040
Db 4981 AAAGAAATCCGAGAAAGCCCACTGGGAATAGCAATGGAATAAACCAGCCGTCCGCGC 5040
Qy 5041 ACTGTGAAAGCTAATAGTAGAGCCGAGCCAGCCCTCGTCTCATCATACCTGTGCGA 5100
Db 5041 ACTGTGAAAGCTAATAGTAGAGCCGAGCCAGCCCTCGTCTCATCATACCTGTGCGA 5100
Qy 5101 AACCCGACAATTTGTTGCACTATAGCGCGAATAGGCTTTTCCAGAGCTCTGTCTCC 5160
Db 5101 AACCCGACAATTTGTTGCACTATAGCGCGAATAGGCTTTTCCAGAGCTCTGTCTCC 5160
Qy 5161 GGTATAGGTCATTTTGAACACCCCTCTCCACTGGGCGAGGCTAATTAATTTTCC 5220
Db 5161 GGTATAGGTCATTTTGAACACCCCTCTCCACTGGGCGAGGCTAATTAATTTTCC 5220
Qy 5221 TTCTTTGCACTCACTTACTAGCCAGTTTATGTTTGAAGGACCAACCGGTTTGT 5280
Db 5221 TTCTTTGCACTCACTTACTAGCCAGTTTATGTTTGAAGGACCAACCGGTTTGT 5280
Qy 5281 GAAGGTCCTAGAACTCAACATTTTGGGAAGCTTCTAGAGGCTATGATGTTCTTT 5340
Db 5281 GAAGGTCCTAGAACTCAACATTTTGGGAAGCTTCTAGAGGCTATGATGTTCTTT 5340
Qy 5341 TGGACATGATATATTTGTTTTTCTTTTGAATTTGCAACATTTTTCAAATTCAT 5400
Db 5341 TGGACATGATATATTTGTTTTTCTTTTGAATTTGCAACATTTTTCAAATTCAT 5400
Qy 5401 GATTTTGTGAACCTTGTGATTTTGAATCCGATTTTTCCTAATTCCTGTGTTT 5460
Db 5401 GATTTTGTGAACCTTGTGATTTTGAATCCGATTTTTCCTAATTCCTGTGTTT 5460
Qy 5461 GAAAAAACTGTGACTTTTCCGAATTAATGAACATTTGTTGCAAGATGATGATCT 5520
Db 5461 GAAAAAACTGTGACTTTTCCGAATTAATGAACATTTGTTGCAAGATGATGATCT 5520
Qy 5521 TTTCAATGAGCGATTTTCTAATAATCCATATTTTCTAATTAAGCTTTCC 5580

```

```

Db 5521 TTTCAATGAGCGATTTTCTAATAATCCATATTTTCTAATTAAGCTTTCC 5580
Qy 5581 TTTTAAATCGAAGTATCTAGATTTGTGAACCTTTTAAATTTCTTAAATTAATGA 5640
Db 5581 TTTTAAATCGAAGTATCTAGATTTGTGAACCTTTTAAATTTCTTAAATTAATGA 5640
Qy 5641 TTTTAAATCGAAGTATCTAGATTTGTGAACCTTTTAAATTTCTTAAATTAATGA 5700
Db 5641 TTTTAAATCGAAGTATCTAGATTTGTGAACCTTTTAAATTTCTTAAATTAATGA 5700
Qy 5701 AACCATTTGATCTGCTGAACAGGCGCTTTGATGATTTGGGTCTTAATTCAGCG 5760
Db 5701 AACCATTTGATCTGCTGAACAGGCGCTTTGATGATTTGGGTCTTAATTCAGCG 5760
Qy 5761 CCCCTGACTTAAGAGAGAAATGCAATTTTAAACACACTAGATGTAATAAATG 5820
Db 5761 CCCCTGACTTAAGAGAGAAATGCAATTTTAAACACACTAGATGTAATAAATG 5820
Qy 5821 AATATCAGATTAATAAAGGAGAAAAAACTATACCTGAAATCCCTGTTTCTAAT 5880
Db 5821 AATATCAGATTAATAAAGGAGAAAAAACTATACCTGAAATCCCTGTTTCTAAT 5880
Qy 5881 ATTTGTTGTGGGAGAACTAATCTGAAGAATCTAATCTAGTTCTCCGCAATAACAATA 5940
Db 5881 ATTTGTTGTGGGAGAACTAATCTGAAGAATCTAATCTAGTTCTCCGCAATAACAATA 5940
Qy 5941 TTATGATTTGGGGAGATATACCTATTAACAGATCAACCAAGAAATGCTCCCAAGAAA 6000
Db 5941 TTATGATTTGGGGAGATATACCTATTAACAGATCAACCAAGAAATGCTCCCAAGAAA 6000
Qy 6001 ACCCAAGAAAGTCTAGAGTTTGTGTTTTCAGAGCCGAAGATAGATAGCATTTCTGA 6060
Db 6001 ACCCAAGAAAGTCTAGAGTTTGTGTTTTCAGAGCCGAAGATAGATAGCATTTCTGA 6060
Qy 6061 ATTAGGTCATCTTTTCCCAAGAGATTTGAAGAAGATAGAAATTTGTGAATTTAGTGGC 6120
Db 6061 ATTAGGTCATCTTTTCCCAAGAGATTTGAAGAAGATAGAAATTTGTGAATTTAGTGGC 6120
Qy 6121 GAGATATCATTTTCTGATTAAGTATGCAATTTTGTGCGGAGAGCCAAACCCGCACTGG 6180
Db 6121 GAGATATCATTTTCTGATTAAGTATGCAATTTTGTGCGGAGAGCCAAACCCGCACTGG 6180
Qy 6181 AGCCGAATTTGAATTAAGTGTGGGTGAGAGTGCAGAGCATGTTTCAAGAG 6240
Db 6181 AGCCGAATTTGAATTAAGTGTGGGTGAGAGTGCAGAGCATGTTTCAAGAG 6240
Qy 6241 TGTAGCAGATGATGTTGCTCTTGTGTGAGGCTGCAATATACATATTAACAA 6300
Db 6241 TGTAGCAGATGATGTTGCTCTTGTGTGAGGCTGCAATATACATATTAACAA 6300
Qy 6301 GAGCATCAGCTGAGCAGCTGCTGCAATGCGTCTTGTGCGCGCCGGAAGCCCGG 6360
Db 6301 GAGCATCAGCTGAGCAGCTGCTGCAATGCGTCTTGTGCGCGCCGGAAGCCCGG 6360
Qy 6361 TTGATGTGCGCAGGCGAGTGGCGAGCGGACCGAGCTATTAAGCAACGCGCAACGCG 6420
Db 6361 TTGATGTGCGCAGGCGAGTGGCGAGCGGACCGAGCTATTAAGCAACGCGCAACGCG 6420
Qy 6421 CGCGGTTCTCAATCCATCCATCCCTTAAGTGAATTAATGAATGATGATTTCCCTG 6480
Db 6421 CGCGGTTCTCAATCCATCCATCCCTTAAGTGAATTAATGAATGATGATTTCCCTG 6480
Qy 6481 CCACTGCTAGTACTCTCTCTGTTTCTGTTGCAATGTTTCAACAGAGCAACGCGCA 6540
Db 6481 CCACTGCTAGTACTCTCTCTGTTTCTGTTGCAATGTTTCAACAGAGCAACGCGCA 6540
Qy 6541 CGGCGAGGCGCGCGCGCGCGCGCAACGAGCAACGAGCAACGAGCAACGAGCAACGAG 6600
Db 6541 CGGCGAGGCGCGCGCGCGCGCGCGCAACGAGCAACGAGCAACGAGCAACGAGCAACGAG 6600
Qy 6601 CGGCAAGAGCAACGAGCAACGAGCGCGCGCGCGAGTGAATTTTCCCGCGGCGCA 6660
Db 6601 CGGCAAGAGCAACGAGCAACGAGCGCGCGCGCGAGTGAATTTTCCCGCGGCGCA 6660

```

6661 GGAACGATCTCTGCGACGACGCGGCGGAGAAACAGATCCGCGGATACGGTACAAAT 6720
6662 GGAACGATCTCTGCGACGACGCGGCGGAGAAACAGATCCGCGGATACGGTACAAAT 6720
6721 CAGCGCGAGCGTGAAGAGACGCGGCGCGCGTGTGCGCGTCCGCGGAGAGCC 6780
6722 CAGCGCGAGCGTGAAGAGACGCGGCGCGCGTGTGCGCGTCCGCGGAGAGCC 6780
6781 GTCCGTGTCTCCCGGCTTCCGCGACGCGCGGTGAGGCGGAGAGAGCGCGCGCT 6840
6782 GTCCGTGTCTCCCGGCTTCCGCGACGCGCGGTGAGGCGGAGAGAGCGCGCGCT 6840
6841 GCGCACGCGCGAGTTCACTGTACGCGCGCGGTGAGGCGGAGAGAGAGAGAG 6900
6842 GCGCACGCGCGAGTTCACTGTACGCGCGCGGTGAGGCGGAGAGAGAGAGAG 6900
6901 ACATTTACAGCTTCACTGTATGTATGCGTGAAGATGAGTGGCGGTTTACTTACGTC 6960
6902 ACATTTACAGCTTCACTGTATGTATGCGTGAAGATGAGTGGCGGTTTACTTACGTC 6960
6961 CCGCGCGCTGTCTTCCCGGCGGCGTCAAAATTTTACCTTCTATAGTACCTTATATA 7020
6962 CCGCGCGCTGTCTTCCCGGCGGCGTCAAAATTTTACCTTCTATAGTACCTTATATA 7020
7021 ACAAACAGCGCGGTGAGAGAGATGTACAGCGCGGTGAGCGCTTACAGCTATCGCGGAC 7080
7022 ACAAACAGCGCGGTGAGAGAGATGTACAGCGCGGTGAGCGCTTACAGCTATCGCGGAC 7080
7081 GACGTCTTCTTCACTGCGCGCGGAGATCTGAGCGATGAGTCAATATCCGCGTGTGCGC 7140
7082 GACGTCTTCTTCACTGCGCGCGGAGATCTGAGCGATGAGTCAATATCCGCGTGTGCGC 7140
7141 CAGACTGCGCGCGGAGATCTGAGCGATGAGTCAATATCCGCGTGTGCGC 7200
7142 CAGACTGCGCGCGGAGATCTGAGCGATGAGTCAATATCCGCGTGTGCGC 7200
7201 GCGGATTTCAACAGCGTGAAGTCCGCGATCTTCACTGAGCGGAGAGAGAGAG 7260
7202 GCGGATTTCAACAGCGTGAAGTCCGCGATCTTCACTGAGCGGAGAGAGAGAG 7260
7261 ATCGACATCGATCGCTGGAATTCATGCGCGAGAGAGAGAGAGAGAGAGAGAG 7320
7262 ATCGACATCGATCGCTGGAATTCATGCGCGAGAGAGAGAGAGAGAGAGAGAG 7320
7321 AACCCAAACATCCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7380
7322 AACCCAAACATCCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7380
7381 CCATGACATCTCTGCGTGTGATGAGAGCGGCTGTGATGAGAGATGAGATGAGT 7440
7382 CCATGACATCTCTGCGTGTGATGAGAGCGGCTGTGATGAGAGATGAGATGAGT 7440
7441 TTGCTATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7500
7442 TTGCTATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7500
7501 TCGGATTTTGGTATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7560
7502 TCGGATTTTGGTATGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7560
7561 TTATCCGATGAGGCGCTTTTGGGACATTTGCGCGGCTTGTGATGATGATGATGAT 7620
7562 TTATCCGATGAGGCGCTTTTGGGACATTTGCGCGGCTTGTGATGATGATGATGAT 7620
7621 AGTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7680
7622 AGTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7680
7681 TTTTGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7740
7682 TTTTGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7740
7740 TTTTGAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7740

7741 GATTAGATTTTGTCTAAATTTGTACTGCTTGTGTTATTCAGATCTTCACTTATAC 7800
7742 GATTAGATTTTGTCTAAATTTGTACTGCTTGTGTTATTCAGATCTTCACTTATAC 7800
7801 GAATTAACCTTAATGATCTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7860
7802 GAATTAACCTTAATGATCTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7860
7861 GCGCTATTTGCTGATGCGCGGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAG 7920
7862 GCGCTATTTGCTGATGCGCGGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAG 7920
7921 AGAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7980
7922 AGAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7980
7981 TTTAGTATTTCCCTTGTGTTGTTGATGATGATGATGATGATGATGATGATGAT 8040
7982 TTTAGTATTTCCCTTGTGTTGTTGATGATGATGATGATGATGATGATGATGAT 8040
8041 AACCTTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8100
8042 AACCTTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8100
8101 TCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8160
8102 TCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8160
8161 GTTAAAGATTTCTTACCTCAACACCTTGTGCGATCTTCAATTTTGTATCTTGA 8220
8162 GTTAAAGATTTCTTACCTCAACACCTTGTGCGATCTTCAATTTTGTATCTTGA 8220
8221 CATTAATGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8280
8222 CATTAATGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8280
8281 TTAGGAAGCTTCTTAAATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 8340
8282 TTAGGAAGCTTCTTAAATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 8340
8341 TGTGCTAATAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8400
8342 TGTGCTAATAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8400
8401 TACGTCCTCAACAG 8460
8402 TACGTCCTCAACAG 8460
8461 CTTTAAAGGTTAATCTGAGATCTCAAGTCAATCAACAACTCAATCAATATATAT 8520
8462 CTTTAAAGGTTAATCTGAGATCTCAAGTCAATCAACAACTCAATCAATATATAT 8520
8521 TATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8580
8522 TATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8580
8581 AAATTAATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8640
8582 AAATTAATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8640
8641 ATTGTAACACTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 8700
8642 ATTGTAACACTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 8700
8701 CTTTCTAATAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8760
8702 CTTTCTAATAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8760
8761 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8820
8762 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8820
8821 TTTATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8880
8822 TTTATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8880


```
Db 8821 TTTATGTCAGAGTGAATGATATGCGCATTTTAAAGAAACATATGATATATAT 8880
Qy ATCTCTGTATATCTAAACAATCTTCTCCATCCCTAAATTAATGCTTACCTTACG 8940
Db 8881 ATCTCTGTATATCTAAACAATCTTCTCCATCCCTAAATTAATGCTTACCTTACG 8940
Qy ACAATTTATATCTAGATCTAGTACAAAGTTGAAACAGTTATTTTGGGACAGAGGAGTAG 9000
Db ACAATTTATATCTAGATCTAGTACAAAGTTGAAACAGTTATTTTGGGACAGAGGAGTAG 9000
Qy TATATATGTTGTGAGACATTAAGGTTATGTTGACTGATATATGCTTCTTAAATGTGAA 9060
Db TATATATGTTGTGAGACATTAAGGTTATGTTGACTGATATATGCTTCTTAAATGTGAA 9060
Qy CATGTCCTTATATGTTTGTATGTTATGAGAGTCTTATCACTTTCCGAGTAC 9120
Db CATGTCCTTATATGTTTGTATGTTATGAGAGTCTTATCACTTTCCGAGTAC 9120
Qy ACATTAATGATTACCATATCATTTGTCAGAAATGATATACCATTTGAAATATCTTCT 9180
Db ACATTAATGATTACCATATCATTTGTCAGAAATGATATACCATTTGAAATATCTTCT 9180
Qy TTTTATGCAAGACTAGATGCAATGTCATTTCTTCTGATCTTATGTCCTTTTCTTCT 9240
Db TTTTATGCAAGACTAGATGCAATGTCATTTCTTCTGATCTTATGTCCTTTTCTTCT 9240
Qy TCGTACATGTTTGCGCTCTCACAAAATTAATATCCAGACATGTTCCAAATAT 9300
Db TCGTACATGTTTGCGCTCTCACAAAATTAATATCCAGACATGTTCCAAATAT 9300
Qy ATTAATATATTTGAGTGTGTTTTCACCACTTATATCTTCTTCTTCTTCTTCTTCT 9360
Db ATTAATATATTTGAGTGTGTTTTCACCACTTATATCTTCTTCTTCTTCTTCTTCT 9360
Qy GTATATATGTTTAACTCTAACAAAACCTTATATATGTTTCTCTTCTTCTTCTTCT 9420
Db GTATATATGTTTAACTCTAACAAAACCTTATATATGTTTCTCTTCTTCTTCTTCT 9420
Qy TCTTGAATGGAATGAGTGGTGGGTTTAACTTTGCGGCTGCTTCTTCTTCTTCTTCT 9480
Db TCTTGAATGGAATGAGTGGTGGGTTTAACTTTGCGGCTGCTTCTTCTTCTTCTTCT 9480
Qy TCTTGAATGGAATGAGTGGTGGGTTTAACTTTGCGGCTGCTTCTTCTTCTTCTTCT 9480
Db TCTTGAATGGAATGAGTGGTGGGTTTAACTTTGCGGCTGCTTCTTCTTCTTCTTCT 9480
Qy TGAATCTGAAAGGGGTCAATCATCTGTCAAGAGAACAGAAAGAAATCTTATATATG 9540
Db TGAATCTGAAAGGGGTCAATCATCTGTCAAGAGAACAGAAAGAAATCTTATATATG 9540
Qy TTTGTTAGTTGTACACACCCCTAGTTGTACATCTGACGAGCTGTAATCATTTCTAGT 9600
Db TTTGTTAGTTGTACACACCCCTAGTTGTACATCTGACGAGCTGTAATCATTTCTAGT 9600
Qy ATCCCAATTTATATATTTCAATTAACATATGTAATGTTCTGTAGTGTGCAAGT 9660
Db ATCCCAATTTATATATTTCAATTAACATATGTAATGTTCTGTAGTGTGCAAGT 9660
Qy CATGTAATCTTATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 9720
Db CATGTAATCTTATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 9720
Qy GGTGTGTACACATGATGATGATGAGGCGAGGTTTGTGACACCTTTTCAACAAAAT 9780
Db GGTGTGTGTACACATGATGATGATGAGGCGAGGTTTGTGACACCTTTTCAACAAAAT 9780
Qy CTATATCTTTCAACAAATGAACCTTGTAGTCCCTTGTAGTAGAATCAATCTCTCT 9840
Db CTATATCTTTCAACAAATGAACCTTGTAGTCCCTTGTAGTAGAATCAATCTCTCT 9840
Qy GAATATGCTATGTTTCAACATGATGATGATGATGATGATGATGATGATGATGAT 9900
Db GAATATGCTATGTTTCAACATGATGATGATGATGATGATGATGATGATGATGAT 9900
Qy CATGTCAAAGTTTATATGTTTATATTTCAATATGAGAACTTTGATATCTTCTAGAC 9960
Db CATGTCAAAGTTTATATGTTTATATTTCAATATGAGAACTTTGATATCTTCTAGAC 9960

Db 9901 CATGTCAAAGTTTATATGTTTATATTTCAATATGAGAACTTTGATATCTTCTAGAC 9960
Qy ACAATCTCTGTAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 10020
Db ACAATCTCTGTAATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 10020
Qy GGGGGGGGGTGAATAGCGGTTTATTAACAATGTATATTTGAGAAATATCTTATGTGA 10080
Db GGGGGGGGGTGAATAGCGGTTTATTAACAATGTATATTTGAGAAATATCTTATGTGA 10080
Qy ATTAATCTAGGTGATATTTTTCATTAAGAGGTCCTTTTATGACTCAATATGTAACA 10140
Db ATTAATCTAGGTGATATTTTTCATTAAGAGGTCCTTTTATGACTCAATATGTAACA 10140
Qy TCAAGGATACATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10200
Db TCAAGGATACATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10200
Qy CCAACACACACACACACACACACATTCACACACACAAATCATGTCAGAGAGCGAA 10260
Db CCAACACACACACACACACACACATTCACACACACAAATCATGTCAGAGAGCGAA 10260
Qy GTCATACAGATCAAAATCTATGCTGAGCGGAGGAGAAATGAAAATCATGAAATGTA 10320
Db GTCATACAGATCAAAATCTATGCTGAGCGGAGGAGAAATGAAAATCATGAAATGTA 10320
Qy AAAACGTCATGACAAATCATGACATGACATGACATGACATGACATGACATGACATGAC 10380
Db AAAACGTCATGACAAATCATGACATGACATGACATGACATGACATGACATGACATGAC 10380
Qy CTGCGAGAAAAGTTCTATTAATACTGCGGCTTGCAGAGAGAAAGAGTGTGC 10440
Db CTGCGAGAAAAGTTCTATTAATACTGCGGCTTGCAGAGAGAAAGAGTGTGC 10440
Qy CATCATCGATCCACACATTAAGTATCTGCGGTTTCACTGTAATCAATCCG 10500
Db CATCATCGATCCACACATTAAGTATCTGCGGTTTCACTGTAATCAATCCG 10500
Qy AGCAATCTCGAATATGTTTATTAAGGTTTCAAGATTCAAAATGTCACAAATCATGAG 10560
Db AGCAATCTCGAATATGTTTATTAAGGTTTCAAGATTCAAAATGTCACAAATCATGAG 10560
Qy TTAATGACAAATTAAGACACACCTAGATTTTATCCAAAGCTGAGCGGTACTTGA 10620
Db TTAATGACAAATTAAGACACACCTAGATTTTATCCAAAGCTGAGCGGTACTTGA 10620
Qy TTAATGACAAATTAAGACACACCTAGATTTTATCCAAAGCTGAGCGGTACTTGA 10620
Db TTAATGACAAATTAAGACACACCTAGATTTTATCCAAAGCTGAGCGGTACTTGA 10620
Qy AGTACCATCAATTTGAATGATATCCCACTTGCATCAATCAATAGTTGATGATGACG 10680
Db AGTACCATCAATTTGAATGATATCCCACTTGCATCAATCAATAGTTGATGATGACG 10680
Qy AGTACCATCAATTTGAATGATATCCCACTTGCATCAATCAATAGTTGATGATGACG 10680
Db AGTACCATCAATTTGAATGATATCCCACTTGCATCAATCAATAGTTGATGATGACG 10680
Qy GTCCATATGAGCGAGTATGAGCATGAGCGGCAATGTAAGTTTAAAGTGAAGTGAAGA 10740
Db GTCCATATGAGCGAGTATGAGCATGAGCGGCAATGTAAGTTTAAAGTGAAGTGAAGA 10740
Qy GTCCATATGAGCGAGTATGAGCATGAGCGGCAATGTAAGTTTAAAGTGAAGTGAAGA 10740
Db GTCCATATGAGCGAGTATGAGCATGAGCGGCAATGTAAGTTTAAAGTGAAGTGAAGA 10740
Qy GCCTGTGCGCCACTGAGAAAGTGTGATGAGGAGGAAAGTATGAGCTTCCACCAAC 10800
Db GCCTGTGCGCCACTGAGAAAGTGTGATGAGGAGGAAAGTATGAGCTTCCACCAAC 10800
Qy GCCTGTGCGCCACTGAGAAAGTGTGATGAGGAGGAAAGTATGAGCTTCCACCAAC 10800
Db GCCTGTGCGCCACTGAGAAAGTGTGATGAGGAGGAAAGTATGAGCTTCCACCAAC 10800
Qy ATCCCAAGTTTGAACAAATCTTGAAGCCCTTACCAATCAATCAACAAATGTAAGTAC 10860
Db ATCCCAAGTTTGAACAAATCTTGAAGCCCTTACCAATCAATCAACAAATGTAAGTAC 10860
Qy TATCTGTATCAAGACCAATCATGATCCGTCTGATCTGCTTATCTTGTATGATGATAC 10920
Db TATCTGTATCAAGACCAATCATGATCCGTCTGATCTGCTTATCTTGTATGATGATAC 10920
Qy TACCTGCAAAAGGCGGATTCAGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 10966
Db TACCTGCAAAAGGCGGATTCAGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 10966
Qy TACCTGCAAAAGGCGGATTCAGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 10966
Db TACCTGCAAAAGGCGGATTCAGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 10966
```

RESULT 2
BD012010

LOCUS	BD012010	10966 bp	DNA	linear	PAT 02-AUG-2002
DEFINITION	Creation of iron-deficiency resisting rice plant.				
ACCESSION	BD012010				
VERSION	BD012010.1	GI:22092199			
KEYWORDS	WO 0101762-A/1.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 10966)				
AUTHORS	Mori,S., Nakanishi,H., Takahashi,M. and Nishizawa,N.				
TITLE	Creation of iron-deficiency resisting rice plant				
JOURNAL	Patent: WO 0101762-A 1 11-JAN-2001; JAPAN SCIENCE AND TECHNOLOGY CORP,SATOSHI MORI,HIROMI NAKANISHI, MICHIO TAKAHASHI,NAOKO NISHIZAWA, OS Hordeum vulgare L. var. Igri PN WO 0101762-A/1 PD 11-JAN-2001 PF 04-JUL-2000 WO 2000JP004425 PR 05-JUL-1999 JP 99P 190318 PI SATOSHI MORI,HIROMI NAKANISHI,MICHIO TAKAHASHI, PI NAOKO NISHIZAWA PC A01H5/00,C12N5/14,C12N15/52 CC				
FEATURES	FH Key Location/Qualifiers.				
source	1..10966 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"				
ORIGIN	Query Match 100.0%; Score 10966; DB 6; Length 10966; Best Local Similarity 100.0%; Pred. No. 0; Matches 10966; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	CTCATCCCATTTGCATATGATATGATTAGCTATCAACGAAAGAAAGATGCGATGCGC	60		
DB	1	CTCATCCCATTTGCATATGATATGATTAGCTATCAACGAAAGAAAGATGCGATGCGC	60		
QY	61	CTGTGTGCATCCCTCACTGGCTTGGGGAATGGGATACCGAGTTTGGTAGAGTATTTT	120		
DB	61	CTGTGTGCATCCCTCACTGGCTTGGGGAATGGGATACCGAGTTTGGTAGAGTATTTT	120		
QY	121	TTAGCATGATGTCGCCGCCACTGCGCAAGAAACTGCGTCAAGCGACTGCAAGAGTT	180		
DB	121	TTAGCATGATGTCGCCGCCACTGCGCAAGAAACTGCGTCAAGCGACTGCAAGAGTT	180		
QY	181	GAGGATGTCATCTTTGTGATGATGACGGAGCTGATGGGTTCCTAACTGAACCCAACTA	240		
DB	181	GAGGATGTCATCTTTGTGATGATGACGGAGCTGATGGGTTCCTAACTGAACCCAACTA	240		
QY	241	GCAATGGGTGATGATGATGATGACGGAGCTGATGGGTTCCTAACTGAACCCAACTA	300		
DB	241	GCAATGGGTGATGATGATGATGACGGAGCTGATGGGTTCCTAACTGAACCCAACTA	300		
QY	301	TTCTCCCAATTTAATAAGGGATATGATGATGATGATGATGATGATGATGATGATGATG	360		
DB	301	TTCTCCCAATTTAATAAGGGATATGATGATGATGATGATGATGATGATGATGATGATG	360		
QY	361	ATTAGAACCAAGCGAGGGCGCGAATAATGCGCGGATTAATGGGGAGGCGCGCAAGCG	420		
DB	361	ATTAGAACCAAGCGAGGGCGCGAATAATGCGCGGATTAATGGGGAGGCGCGCAAGCG	420		
QY	421	CACGCGCGTCACTGGCAAGCGCGTGCAGAGCGTGCAGAGCGTGCAGCGTGCCTGCT	480		
DB	421	CACGCGCGTCACTGGCAAGCGCGTGCAGAGCGTGCAGAGCGTGCAGCGTGCCTGCT	480		
QY	481	TCTTGTGCGCGGGCGCGGTGCTGTGTGCGTCAAGCAACGGCTATATAGGACGCTAATC	540		
DB	481	TCTTGTGCGCGGGCGCGGTGCTGTGTGCGTCAAGCAACGGCTATATAGGACGCTAATC	540		
QY	541	ACCGTACTCAATCCGTCGCCCACTGCTTCTTATACCGCTACTGATGATTTCTGTGT	600		

[illegible]

QY 1681 GCATTTGCACTCATCCCGCAAGAGGGTGGAGATGACATGCACTGCTGGAATCCAT 1740
DB 1681 GCATTTGCACTCATCCCGCAAGAGGGTGGAGATGACATGCACTGCTGGAATCCAT 1740
QY 1741 CGCGGACAAAGAACACACCGGCAATGTCATCATTAACCCCAACAACCCGTGGGAGGT 1800
DB 1741 CGCGGACAAAGAACACACCGGCAATGTCATCATTAACCCCAACAACCCGTGGGAGGT 1800
QY 1801 TTACTCCAGACCACTGTGCAAGGTTTCAATCCTTTGGCTTGGCTGGAATGGA TTCA 1860
DB 1801 TTACTCCAGACCACTGTGCAAGGTTTCAATCCTTTGGCTTGGCTGGAATGGAATGGA 1860
QY 1861 GTTCAGTGCACCTGCTGAATTTCTTTTGGCAATGCAATGCAATGCTGATGTTGCTCAATTA 1920
DB 1861 GTTCAGTGCACCTGCTGAATTTCTTTTGGCAATGCAATGCAATGCTGATGTTGCTCAATTA 1920
QY 1921 GGTGGCGGAGGTGGCGAAAGAGCTCGGAATATTTGGTGAATGCTGACGAGATTAACGGCAA 1980
DB 1921 GGTGGCGGAGGTGGCGAAAGAGCTCGGAATATTTGGTGAATGCTGACGAGATTAACGGCAA 1980
QY 1981 GCTGGTTCTGGGACGCGCCCGTTCATCCCAATGGAGTGTGGGGCAATCACCCCTGT 2040
DB 1981 GCTGGTTCTGGGACGCGCCCGTTCATCCCAATGGAGTGTGGGGCAATCACCCCTGT 2040
QY 2041 GCTGTCCATAGAGGTCTCTGTCGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2100
DB 2041 GCTGTCCATAGAGGTCTCTGTCGAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2100
QY 2101 AGCGGTGTAACGACCCCGAAGAAATCTTACAGAAACCTTAACTTAAATCTTAAATCTTAAAT 2160
DB 2101 AGCGGTGTAACGACCCCGAAGAAATCTTACAGAAACCTTAACTTAAATCTTAAATCTTAAAT 2160
QY 2161 TTCTTTTCAAAATGCTACTAAGGTGATTAATAGTACTGCTGATCAATTAATTTGCTAAAT 2220
DB 2161 TTCTTTTCAAAATGCTACTAAGGTGATTAATAGTACTGCTGATCAATTAATTTGCTAAAT 2220
QY 2221 TTGTACTGCAATTTTGTGTAGATCTCTACATCAATTAACGAATTAACGAATTAACGAAT 2280
DB 2221 TTGTACTGCAATTTTGTGTAGATCTCTACATCAATTAACGAATTAACGAATTAACGAAT 2280
QY 2281 CAGACCCAGAACCTTCAATCAGGTCACTCTTGTGATTTTACCTGCTTCAAGAAATTA 2340
DB 2281 CAGACCCAGAACCTTCAATCAGGTCACTCTTGTGATTTTACCTGCTTCAAGAAATTA 2340
QY 2341 GTCTTTGGTATTTACTCTCTCTTGTCTCAATTTTGTCTCGGCTCTTAATGTTGAGGCTGCC 2400
DB 2341 GTCTTTGGTATTTACTCTCTCTTGTCTCAATTTTGTCTCGGCTCTTAATGTTGAGGCTGCC 2400
QY 2401 CAGGTGCAATGTCAGATGACGTTTTTCACTTAAGTTTGAAGTCAAGTCAAGTCAAGT 2460
DB 2401 CAGGTGCAATGTCAGATGACGTTTTTCACTTAAGTTTGAAGTCAAGTCAAGTCAAGT 2460
QY 2461 ACACCTGAGTATTTTAACTTTTGTGCTTGTGATCCGATTAATTAATTAATTAATTAATTA 2520
DB 2461 ACACCTGAGTATTTTAACTTTTGTGCTTGTGATCCGATTAATTAATTAATTAATTAATTA 2520
QY 2521 CTGAACCTACTGTAATATTAACACATGTTCTTCAAGATATTAATGATGCTATGAGG 2580
DB 2521 CTGAACCTACTGTAATATTAACACATGTTCTTCAAGATATTAATGATGCTATGAGG 2580
QY 2581 CATGCCATATTTCTTTGGGTCAAGATGACATGATGGAACCTCTTTTGAAGAAATGAT 2640
DB 2581 CATGCCATATTTCTTTGGGTCAAGATGACATGATGGAACCTCTTTTGAAGAAATGAT 2640
QY 2641 ACATGTAATGATGATTAACATTTTAAAGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 2700
DB 2641 ACATGTAATGATGATTAACATTTTAAAGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 2700
QY 2701 ATTGTTCTCTGTAATTAACAGATTTTAACTGTTGTTTGTGATCTAATCAAAATTTGTTGAG 2760
DB 2701 ATTGTTCTCTGTAATTAACAGATTTTAACTGTTGTTTGTGATCTAATCAAAATTTGTTGAG 2760

QY 2761 GCAGCTCTTCTCAGATTTCTTGAGAACACAAAGAAATTTCTTTAAGGCAATTTAGT 2820
DB 2761 GCAGCTCTTCTCAGATTTCTTGAGAACACAAAGAAATTTCTTTAAGGCAATTTAGT 2820
QY 2821 CTGCTAAAGAAATCATGAGATATGCTACAAACAAATTAAGGAAACAAATATCATTA 2880
DB 2821 CTGCTAAAGAAATCATGAGATATGCTACAAACAAATTAAGGAAACAAATATCATTA 2880
QY 2881 TGTCTTCAAGGCAAGAGATCAATGTTGTATGTTAGGCTTATTTTGAAGTAAA 2940
DB 2881 TGTCTTCAAGGCAAGAGATCAATGTTGTATGTTAGGCTTATTTTGAAGTAAA 2940
QY 2941 AAATCTTGAAGAGTGTCAATATCATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000
DB 2941 AAATCTTGAAGAGTGTCAATATCATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000
QY 3001 ATGATTCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3060
DB 3001 ATGATTCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3060
QY 3061 TGAAGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3120
DB 3061 TGAAGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3120
QY 3121 CTGATATCTCTTGTCCATCTTTTGTATTAATGATGATTAATGATTAATGATTAATTA 3180
DB 3121 CTGATATCTCTTGTCCATCTTTTGTATTAATGATGATTAATGATTAATGATTAATTA 3180
QY 3181 ACAAAATATCCATGCTAATTTTGAAGAAATTAATTAATTTTCTGTAACCAATGAACTCC 3240
DB 3181 ACAAAATATCCATGCTAATTTTGAAGAAATTAATTAATTTTCTGTAACCAATGAACTCC 3240
QY 3241 GTATTAATCTTCCATGTAACAGGTGAAACCTGAACTTAACTTTTGAAGAAATTAAGCA 3300
DB 3241 GTATTAATCTTCCATGTAACAGGTGAAACCTGAACTTAACTTTTGAAGAAATTAAGCA 3300
QY 3301 TGAATGATTTTGTCTGCAAGCTCGCAAAAGAAATGAAATCTTAATGCTTGTGAGT 3360
DB 3301 TGAATGATTTTGTCTGCAAGCTCGCAAAAGAAATGAAATCTTAATGCTTGTGAGT 3360
QY 3361 GAATCCATGTTGATTTTGAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 3420
DB 3361 GAATCCATGTTGATTTTGAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 3420
QY 3421 CATATAATGATTAACATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 3480
DB 3421 CATATAATGATTAACATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 3480
QY 3481 AAGACTTTTAAACAATATCTGTGAAGACTGATATCATGATCTTCTCTTGTGTTTC 3540
DB 3481 AAGACTTTTAAACAATATCTGTGAAGACTGATATCATGATCTTCTCTTGTGTTTC 3540
QY 3541 TTGAACGATGATCTTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
DB 3541 TTGAACGATGATCTTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
QY 3601 TCAAAATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660
DB 3601 TCAAAATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660
QY 3661 TTTTGAATGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3720
DB 3661 TTTTGAATGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3720
QY 3721 CCTCTAAATGTAAGGAGTGTCTTGAATGGAACCTGGGCTGGATTAATCTTTTGTGCTG 3780
DB 3721 CCTCTAAATGTAAGGAGTGTCTTGAATGGAACCTGGGCTGGATTAATCTTTTGTGCTG 3780
QY 3781 TGTTCATCTTCTTCAAGATGCTCGAAGAGTCAATCATCTGTCGTAAGAGAACAA 3840
DB 3781 TGTTCATCTTCTTCAAGATGCTCGAAGAGTCAATCATCTGTCGTAAGAGAACAA 3840
QY 3841 GAAAGAAATTCGAGGATGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 3900

```
Db 3841 GAAAGAAATTCGAGCGATGATTCCTAGTGTATATCTGACGTGAAGCTGTAATCATTTCC 3900
Qy 3901 CAGTATCCCCATCTATATCTTTCAATAAATGAACTTTTATGTTCTCTATGATAGAAAT 3960
Db 3901 CAGTATCCCCATCTATATCTTTCAATAAATGAACTTTTATGTTCTCTATGATAGAAAT 3960
Qy 3961 CAACATCTCCTGTAATATGTTCTGTTGTTGTTGCTGGAAGAAACATAGGAAATGTTAT 4020
Db 3961 CAACATCTCCTGTAATATGTTCTGTTGTTGTTGCTGGAAGAAACATAGGAAATGTTAT 4020
Qy 4021 GTTAGTAAAGTTACATGGCGTCGAAGATCTTGAAGTTTTTTTTTTTTTGGGGGGGG 4080
Db 4021 GTTAGTAAAGTTACATGGCGTCGAAGATCTTGAAGTTTTTTTTTTTTTGGGGGGGG 4080
Qy 4081 GGGGGGGGGGCTTTGATATATCTTTAGTACAGTTCTCTCAAGTTATGTAAGCA 4140
Db 4081 GGGGGGGGGGCTTTGATATATCTTTAGTACAGTTCTCTCAAGTTATGTAAGCA 4140
Qy 4141 CTTTGTAAACAAATTGATGATTTGGATCATGATAGTAATTAAGTACAGTACCTTGGT 4200
Db 4141 CTTTGTAAACAAATTGATGATTTGGATCATGATAGTAATTAAGTACAGTACCTTGGT 4200
Qy 4201 AAGCAAAACCTTACCTATGTTAGCTCACTAAGGTGGCTTTGGTTGAGAGAGAGAA 4260
Db 4201 AAGCAAAACCTTACCTATGTTAGCTCACTAAGGTGGCTTTGGTTGAGAGAGAGAA 4260
Qy 4261 GGAATCAATTGATATATCCCAATCATGAGTAATCATGTTGTTGCTACCACTTTT 4320
Db 4261 GGAATCAATTGATATATCCCAATCATGAGTAATCATGTTGTTGCTACCACTTTT 4320
Qy 4321 CTACAACTCTAGTACGTCAGTGGCTGACGTAACACACACGTCGCAACATAT 4380
Db 4321 CTACAACTCTAGTACGTCAGTGGCTGACGTAACACACACGTCGCAACATAT 4380
Qy 4381 CTCTGTCAAAATCGGACCCCAAGATTAATCTCAAGCTGGAAGCAACCAATTTG 4440
Db 4381 CTCTGTCAAAATCGGACCCCAAGATTAATCTCAAGCTGGAAGCAACCAATTTG 4440
Qy 4441 AAGAGAGAAACCTTCAACAAAGACCTTGAAGTCCCGCCCAATGATGTTAGCCGCG 4500
Db 4441 AAGAGAGAAACCTTCAACAAAGACCTTGAAGTCCCGCCCAATGATGTTAGCCGCG 4500
Qy 4501 TCGGAGCCCGAGTGCACATGCGGACCAACCACTCCACGCGGGGAGACGTCAC 4560
Db 4501 TCGGAGCCCGAGTGCACATGCGGACCAACCACTCCACGCGGGGAGACGTCAC 4560
Qy 4561 CGATTACTGAAATTCGCCAAACATCTTAAATTTGTGAACAAATTTAAACACAGAA 4620
Db 4561 CGATTACTGAAATTCGCCAAACATCTTAAATTTGTGAACAAATTTAAACACAGAA 4620
Qy 4621 ATTTTGAATTTGTGAACAAATTTTAAACGCGTATCTGAAACATTTTCAAAATTTG 4680
Db 4621 ATTTTGAATTTGTGAACAAATTTTAAACGCGTATCTGAAACATTTTCAAAATTTG 4680
Qy 4681 GATCAAAATTTTAAAGCACTTCTTCAATTTGACAAATTTTAAATTTAAATTA 4740
Db 4681 GATCAAAATTTTAAAGCACTTCTTCAATTTGACAAATTTTAAATTTAAATTA 4740
Qy 4741 GTTCAACAAATTTTGAATTTTAAATTTAGCGAAACATTTGAAATTTCTAAATTTT 4800
Db 4741 GTTCAACAAATTTTGAATTTTAAATTTAGCGAAACATTTTGAATTTCTAAATTTT 4800
Qy 4801 CGAAATTTGGAACATTTTCTATTTCTGAACAAATTTGAATTTGAATTTGGA 4860
Db 4801 CGAAATTTGGAACATTTTCTATTTCTGAACAAATTTGAATTTGAATTTGGA 4860
Qy 4861 ATAAATTTTGGAAATGCGATTTTGTGAAATTTCTGAACATTTTGAATTTGA 4920
Db 4861 ATAAATTTTGGAAATGCGATTTTGTGAAATTTCTGAACATTTTGAATTTGA 4920
Qy 4921 CTTTAAAGGTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4980
Db 4921 CTTTAAAGGTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4980

Db 4921 CTTTAAAGGTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4980
Qy 4981 AAGAAATTCGAGAAACCACTGGGAATGCAATGAAACCAAGCGCTCCGCGC 5040
Db 4981 AAGAAATTCGAGAAACCACTGGGAATGCAATGAAACCAAGCGCTCCGCGC 5040
Qy 5041 ACTGTAAAGCTTAATGAGCGCGCCCAAGCTCTGCTCATCATATACCTGTGGA 5100
Db 5041 ACTGTAAAGCTTAATGAGCGCGCCCAAGCTCTGCTCATCATATACCTGTGGA 5100
Qy 5101 AACCCGACAAATTCGTTGACATATGCGGCAATATAGGCTTTTCCAGAGCTCTCTTCC 5160
Db 5101 AACCCGACAAATTCGTTGACATATGCGGCAATATAGGCTTTTCCAGAGCTCTCTTCC 5160
Qy 5161 GATTATGGGCTATTGCAACCCCTCTCCACTTGGGCGAGCTATATATTTTTC 5220
Db 5161 GATTATGGGCTATTGCAACCCCTCTCCACTTGGGCGAGCTATATATTTTTC 5220
Qy 5221 TTTTTCGACCTCACTTACACGCGAGTTTATGTTTGAAGGCAACCGGTTTTGT 5280
Db 5221 TTTTTCGACCTCACTTACACGCGAGTTTATGTTTGAAGGCAACCGGTTTTGT 5280
Qy 5281 GAAAGTTCTAGAACTCAACCAATTTTGGAAAGCTTGAAGGCTATGAAATGTTCTT 5340
Db 5281 GAAAGTTCTAGAACTCAACCAATTTTGGAAAGCTTGAAGGCTATGAAATGTTCTT 5340
Qy 5341 TGGACATGATATATTTGTTTCTTTTCAATTTGCAATCTTTTCAATTCAT 5400
Db 5341 TGGACATGATATATTTGTTTCTTTTCAATTTGCAATCTTTTCAATTCAT 5400
Qy 5401 GATTTTGTGAACTTGTGATTTTGAATCCGTAATTTTCTTAAATCCGTTTT 5460
Db 5401 GATTTTGTGAACTTGTGATTTTGAATCCGTAATTTTCTTAAATCCGTTTT 5460
Qy 5461 GAAAAAATCTGTGACCTTTTCCGAAATTAATGAACATTTGTTGCAAGATGATCT 5520
Db 5461 GAAAAAATCTGTGACCTTTTCCGAAATTAATGAACATTTGTTGCAAGATGATCT 5520
Qy 5521 TTTCAATGAGCGATTTTCTTAAATATCCACATTTTCTTCAATTAACCTTCC 5580
Db 5521 TTTCAATGAGCGATTTTCTTAAATATCCACATTTTCTTCAATTAACCTTCC 5580
Qy 5581 TTTTAACTGAACTATCTTGAATTTGTTGTTTAAATTTTAAATTTTAAATTA 5640
Db 5581 TTTTAACTGAACTATCTTGAATTTGTTGTTTAAATTTTAAATTTTAAATTA 5640
Qy 5641 TTTTAACTGAACTATCTTGAATTTGTTGTTTAAATTTTAAATTTTAAATTA 5700
Db 5641 TTTTAACTGAACTATCTTGAATTTGTTGTTTAAATTTTAAATTTTAAATTA 5700
Qy 5701 AACCATGACTCGTGAACAGGGGCTTCAATATGATTTGGTGTCTTAAATCCAGCG 5760
Db 5701 AACCATGACTCGTGAACAGGGGCTTCAATATGATTTGGTGTCTTAAATCCAGCG 5760
Qy 5761 CCCCTGACTACTAAACGAAGATTTGAATTTTAAACCACTACTATGTTAAATTA 5820
Db 5761 CCCCTGACTACTAAACGAAGATTTGAATTTTAAACCACTACTATGTTAAATTA 5820
Qy 5821 AATATCAAGATTAATAAGGGAATAATAATAATAATAATAATAATAATAATAATA 5880
Db 5821 AATATCAAGATTAATAAGGGAATAATAATAATAATAATAATAATAATAATAATA 5880
Qy 5881 ATTTGTTGTTGGGAGAACTTAACTGAAAGCACTTAACTGTTCTCCGCAATTA 5940
Db 5881 ATTTGTTGTTGGGAGAACTTAACTGAAAGCACTTAACTGTTCTCCGCAATTA 5940
Qy 5941 TTAATGTTGGGGGAGATTAATCTATTAACATCAACCAAGATGCTCTCAAGAAA 6000
Db 5941 TTAATGTTGGGGGAGATTAATCTATTAACATCAACCAAGATGCTCTCAAGAAA 6000
Qy 6001 ACCCAAGAAAGTCTAGAGTTTGTTCAGAGACGAAAGATAGATGATCTGCA 6060
Db 6001 ACCCAAGAAAGTCTAGAGTTTGTTCAGAGACGAAAGATAGATGATGATCTGCA 6060
```

QY	6061	ATYAGTCCATCTTTTCTTCCCAAGATTAAGAAAGAAAGATGAATTCGATTATGATGGCGC	6120
Db	6061	ATTAGGTCCATCTTTTCTTCCCAAGATTAAAGAAAGATGAATTCGATTATGATGGCGC	6120
QY	6121	GAGATATCATTTCTGGATTAGGTAACAAATGTTTTCGCGCAAGCCAAACCCCGCAGTGG	6180
Db	6121	GAGATATCATTTCTGGATTAGGTAACAAATGTTTTCGCGCAAGCCAAACCCCGCAGTGG	6180
QY	6181	AGCCGGAAATTGGAAATTGATGGGTGAGATGAGAAAGCATTGGTTCAATGGTTCTCAAGAG	6240
Db	6181	AGCCGGAAATTGGAAATTGATGGGTGAGATGAGAAAGCATTGGTTCTCAAGAG	6240
QY	6241	TGTAGCCAGTAGTGTGTGCTCTCCGTGCTGAGAGCTGATATACAGTACATAAACA	6300
Db	6241	TGTAGCCAGTAGTGTGTGCTCTCCGTGCTGAGAGCTGATATACAGTACATAAACA	6300
QY	6301	GACGATCATGCTGGCAGCGCTGCTGCATGCGTCTTCTTGTCGCGCCCGGAAACCCCGG	6360
Db	6301	GACGATCATGCTGGCAGCGCTGCTGCATGCGTCTTCTTGTCGCGCCCGGAAACCCCGG	6360
QY	6361	TTGATGTGCGAGCGAGTGGCGAACGGGACCGAGCGGCTATTAACGACGGCCAAACACACGC	6420
Db	6361	TTGATGTGCGAGCGAGTGGCGAACGGGACCGAGCGGCTATTAACGAGCCACGACCGC	6420
QY	6421	CGCGGTTCTCAATCCATCCATCCCTTAGCTGATTGATTTGACTAGCTAGTTCATTTCCCTG	6480
Db	6421	CGCGGTTCTCAATCCATCCATCCCTTAGCTGATTGATTTGACTAGCTAGTTCATTTCCCTG	6480
QY	6481	CCACACTGCTAGTACTCTCTCTCGTTTCTCTGTGCGCAATGTTACACAGACGCAACGGCCA	6540
Db	6481	CCACACTGCTAGTACTCTCTCTCGTTTCTCTGTGCGCAATGTTACACAGACGCAACGGCCA	6540
QY	6541	CGGCGAGGCGCCCGCCCGCGCGCCCAACGCGCAAGCAACGAGACGCGCGCCCGCGGAA	6600
Db	6541	CGGCGAGGCGCCCGCCCGCGCGCCCAACGCGCAAGCAACGAGACGCGCGCCCGCGGAA	6600
QY	6601	CGGCAAGACGCAACGCGGACGCGCGCGCGCGCGCTGGAGTGAATTTGCGCCCGGGCA	6660
Db	6601	CGGCAAGACGCAACGCGGACGCGCGCGCGCGCGCTGGAGTGAATTTGCGCCCGGGCA	6660
QY	6661	GGAAGGACCTCTTGGGAGAGAGAGGGGGGCGAAGAACAGCATCCGGGCGATATGCGTCAAGAT	6720
Db	6661	GGAAGGACCTCTTGGGAGAGAGAGGGGGGCGAAGAACAGCATCCGGGCGATATGCGTCAAGAT	6720
QY	6721	CAGCGCGAGCGTGAAGAGAGAGCGGAGCGCGCGCGCTGTGCGCTGCGCCACGATGAACC	6780
Db	6721	CAGCGCGAGCGTGAAGAGAGAGCGGAGCGCGCGCGCTGTGCGCTGCGCCACGATGAACC	6780
QY	6781	GTCCGTGTCTCCCGGCTTTCGCGCAACGCGCGTGTGAAGCGCGAGCGCGTCCGCGCGCGCT	6840
Db	6781	GTCCGTGTCTCCCGGCTTTCGCGCAACGCGCGTGTGAAGCGCGAGCGCGTCCGCGCGCGCT	6840
QY	6841	GGCGCACGGGACGTTCAACTGCTACGCGCGCGGCGTCCGCGCGCGCGCAAGAGTGA	6900
Db	6841	GGCGCACGGGACGTTCAACTGCTACGCGCGCGGCGTCCGCGCGCGCGCAAGAGTGA	6900
QY	6901	ACATTTACAGCTTACACCGTAATGTATGCGTGAAGCATGACGCGCGGTTTACTTACGTGC	6960
Db	6901	ACATTTACAGCTTACACCGTAATGTATGCGTGAAGCATGACGCGCGGTTTACTTACGTGC	6960
QY	6961	CCGCGCGGCTTCTTCCCGGCGGTGCGTCAAAATTTTACCTTCTATAGTACCTTATTA	7020
Db	6961	CCGCGCGGCTTCTTCCCGGCGGTGCGTCAAAATTTTACCTTCTATAGTACCTTATTA	7020
QY	7021	ACAAACAGCGCGGTAGCAGACACTTGTACAGGCGGTGCCTTACAGATATCGGCGAC	7080
Db	7021	ACAAACAGCGCGGTAGCAGACACTTGTACAGGCGGTGCCTTACAGATATCGGCGAC	7080
QY	7081	GACGCTTTCTCAACCGCGCGGGAATCTCAGGCGATCGAATCATATCCCGGTGTGGCC	7140
Db	7081	GACGCTTTCTCAACCGCGCGGGAATCTCAGGCGATCGAATCATATCCCGGTGTGGCC	7140

QY	7141	CAGATCGCGGGGCAACAATCTAGCTTCCCGGGCCAGGCTATCCAAATTACGAGGCCGCA	7200
Db	7141	CAGACTGCCGGCCACATACACTGCTTCCCGGGCCAGGCTATCCAAATTACGAGGCCGCA	7200
QY	7201	GCGGCAITTCMAACAGCTGAGAGTCCGGGCACTTCGACTCATCCCGGCAAGGGGTGGAG	7260
Db	7201	GCGGCAITTCMAACAGCTGAGAGTCCGGGCACTTCGACTCATCCCGGCAAGGGGTGGAG	7260
QY	7261	ATCGACATCGACTCGCTGGAAATCCAGTCGCGACAAACAACCAACCGGATGATCATATA	7320
Db	7261	ATCGACATCGACTCGCTGGAAATCCAGTCGCGACAAACAACCAACCGGATGATCATATA	7320
QY	7321	AACCCAAACAATCCGATGGGAGCGGTTTACTCTCTACGACCATCTGGCCAAAGTTTTCAT	7380
Db	7321	AACCCAAACAATCCGATGGGAGCGGTTTACTCTCTACGACCATCTGGCCAAAGTTTTCAT	7380
QY	7381	CCATGCATCTCTGCTGCTGCTGATGACCGGTCGTGTTGAACAATGATATATGCAATTCGCT	7440
Db	7381	CCATGCATCTCTGCTGCTGCTGATGACCGGTCGTGTTGAACAATGATATATGCAATTCGCT	7440
QY	7441	TTGCTAATCGTGTGCTGATGATGCTGTTTGGTTATCAGGTCGGGAGAGTGGCAAGAAC	7500
Db	7441	TTGCTAATCGTGTGCTGATGATGCTGTTTGGTTATCAGGTCGGGAGAGTGGCAAGAAC	7500
QY	7501	TCGGAATATGCTGATTCGCTGACGAGGTTTACGGCAAACTGTTCTTGCGAGCGCCCGT	7560
Db	7501	TCGGAATATGCTGATTCGCTGACGAGGTTTACGGCAAACTGTTCTTGCGAGCGCCCGT	7560
QY	7561	TTATCCCGATGGGCGCTCTTTGGGGCAATGAGCCCGGCTCTGTGCAATGATATCTCTGCCA	7620
Db	7561	TTATCCCGATGGGCGCTCTTTGGGGCAATGAGCCCGGCTCTGTGCAATGATATCTCTGCCA	7620
QY	7621	AGTCGTGATAGTGCCTGATGCGCACTTGGATGATGGTGGCGGTGACGACCCCAACAAGA	7680
Db	7621	AGTCGTGATAGTGCCTGATGCGCACTTGGATGATGGTGGCGGTGACGACCCCAACAAGA	7680
QY	7681	TTTTAGAAAACTAAGTAGCTTTAGCTCCCTATCATTTCTCATATGCTACTGTGGG	7740
Db	7681	TTTTAGAAAACTAAGTAGCTTTAGCTCCCTATCATTTCTCATATGCTACTGTGGG	7740
QY	7741	GATTAGATTTTGGCTAATTTGTACTGCTTGGTTATTCAGATCTCTACGCTATTTAC	7800
Db	7741	GATTAGATTTTGGCTAATTTGTACTGCTTGGTTATTCAGATCTCTACGCTATTTAC	7800
QY	7801	GAATTAACCTTAATGTCTCAACGGACCCAGCAACTTCGTTCAAGTTAGTCTTGGTCTT	7860
Db	7801	GAATTAACCTTAATGTCTCAACGGACCCAGCAACTTCGTTCAAGTTAGTCTTGGTCTT	7860
QY	7861	GCCCATTTTGCTCATGTCGCCGTGTGGATGATGTCAAATGACCGGCTTCAAGTTAGTAT	7920
Db	7861	GCCCATTTTGCTCATGTCGCCGTGTGGATGATGTCAAATGATGACCGGCTTCAAGTTAGTAT	7920
QY	7921	AGAGTTTGTTAAGTGTGAATGTGCAAGTCCMAACATGATGAGAGAAAGTACATCTATT	7980
Db	7921	AGAGTTTGTTAAGTGTGAATGTGCAAGTCCMAACATGATGAGAGAAAGTACATCTATT	7980
QY	7981	TTTAGTACTATCCCTTGTGTTGTTGATTTCCATAAATTAATAACACAAAGCCAGAAC	8040
Db	7981	TTTAGTACTATCCCTTGTGTTGTTGATTTCCATAAATTAATAACACAAAGCCAGAAC	8040
QY	8041	AACATTTGAATAGAACTATTTTCTTAGAAAATATACATGTGATTTTGAAGCAGCCATAT	8100
Db	8041	AACATTTGAATAGAACTATTTTCTTAGAAAATATACATGTGATTTTGAAGCAGCCATAT	8100
QY	8101	TCCTTTGCATCAAGTAGCAATATATTAACCTTGATTTGATCTACGAGTATACCATGTT	8160
Db	8101	TCCTTTGCATCAAGTAGCAATATATTAACCTTGATTTGATCTACGAGTATACCATGTT	8160
QY	8161	GTTAGAGATTTCTTAACTTACACACCGCTGTGCTGATCTTCATATTTTGAATTCCTTGA	8220
Db	8161	GTTAGAGATTTCTTAACTTACACACCGCTGTGCTGATCTTCATATTTTGAATTCCTTGA	8220
QY	8221	CATTATGTTCTCTTATGATTCACACAACCTTAATATGATTTTGTGCTATCCAAATGT	8280

Db 8221 CATTATGTTCTCTTAAGATTACACAACTTAATTAAGATTTTGTGCTACCAATTTGT 8280
Qy 8281 TTAGGAAGCTCTTCTTAATAATCTTGAAGAACACAAAGCAGATTTCTTTAAGAGATAT 8340
Db 8281 TTAGGAAGCTCTTCTTAATAATCTTGAAGAACACAAAGCAGATTTCTTTAAGAGATAT 8340
Qy 8341 TGGCTACTAAGGAATCATAGAGATATGTTATAGGAAATTAAGGAAACAAATATAT 8400
Db 8341 TGGCTACTAAGGAATCATAGAGATATGTTATAGGAAATTAAGGAAACAAATATAT 8400
Qy 8401 TACGTGCTCAACAAGCAGAGAGATGATTTGTATGTAAGCTTAAGATGACTTA 8460
Db 8401 TACGTGCTCAACAAGCAGAGAGATGATTTGTATGTAAGCTTAAGATGACTTA 8460
Qy 8461 CTTTAAAGGTTAACTGGAGATCTCAGTCATCCACAAACAAATCAATCAAAATATAT 8520
Db 8461 CTTTAAAGGTTAACTGGAGATCTCAGTCATCCACAAACAAATCAATCAAAATATAT 8520
Qy 8521 TATGTTTGTATAGATCTTTTGAAGATGATGATTTGAAGAAATTAAGAGAGTGT 8580
Db 8521 TATGTTTGTATAGATCTTTTGAAGATGATGATTTGAAGAAATTAAGAGAGTGT 8580
Qy 8581 AAATTAATTTAGAGATACTTCCGATATCATTTGTTCATTTTGTATTACTGTAA 8640
Db 8581 AAATTAATTTAGAGATACTTCCGATATCATTTGTTCATTTTGTATTACTGTAA 8640
Qy 8641 ATTGGAACACTCAATCATATTAACAAAGATTTCTCCCATTTTGAAGATGACTT 8700
Db 8641 ATTGGAACACTCAATCATATTAACAAAGATTTCTCCCATTTTGAAGATGACTT 8700
Qy 8701 CCTTCTTAATACCATGATTAATTAATCTTCAATGTAACAGGTCAACTTAACATCTTTT 8760
Db 8701 CCTTCTTAATACCATGATTAATTAATCTTCAATGTAACAGGTCAACTTAACATCTTTT 8760
Qy 8761 GAGAGAGATCATACGACATTAATTTTGTGCAAGCTGCAAGAGAGAAATCTGTAT 8820
Db 8761 GAGAGAGATCATACGACATTAATTTTGTGCAAGCTGCAAGAGAGAAATCTGTAT 8820
Qy 8821 TTTATGTCAGATGAGATGATATAGCCATTTTAAAGAAACATATATGGAATATAT 8880
Db 8821 TTTATGTCAGATGAGATGATATAGCCATTTTAAAGAAACATATATGGAATATAT 8880
Qy 8881 ATCTCTGTTATTAATTAATCAATCTTCCATCTTAATTAATGCTTCACTTAC 8940
Db 8881 ATCTCTGTTATTAATTAATCAATCTTCCATCTTAATTAATGCTTCACTTAC 8940
Qy 8941 ACAATTTTATCTAGATCTAGATCAAAAGTTGAACAGTTATTTGGACAGAGGAGTATG 9000
Db 8941 ACAATTTTATCTAGATCTAGATCAAAAGTTGAACAGTTATTTGGACAGAGGAGTATG 9000
Qy 9001 TATATATGTTGTGAGAACATTAAGTTATGTTGAATGATATATGCTTTAAATGTGAA 9060
Db 9001 TATATATGTTGTGAGAACATTAAGTTATGTTGAATGATATATGCTTTAAATGTGAA 9060
Qy 9061 CATGTTCTCTAATGTTTGTATGTTGATGAGAACTTCTATACATTTCCGAGATGACTAC 9120
Db 9061 CATGTTCTCTAATGTTTGTATGTTGATGAGAACTTCTATACATTTCCGAGATGACTAC 9120
Qy 9121 ACAATTAATGATTAATCATATCTGTCAGAAATGATTAACACATTTGAATATCTTTCT 9180
Db 9121 ACAATTAATGATTAATCATATCTGTCAGAAATGATTAACACATTTGAATATCTTTCT 9180
Qy 9181 TTTTATGCAAGAGATGAGATGATGATCTTTCTGATGCTATGATGCTTTTCTTTC 9240
Db 9181 TTTTATGCAAGAGATGAGATGATGATCTTTCTGATGCTATGATGCTTTTCTTTC 9240
Qy 9241 TCGTTACATGTTTGTGCTCTCAAAAATTAATTAACAGCAGATGCTCAAAATGAT 9300
Db 9241 TCGTTACATGTTTGTGCTCTCAAAAATTAATTAACAGCAGATGCTCAAAATGAT 9300
Qy 9301 ATTAAATATTTGAGGTGTTTTCACCACTTATATCTTTCAAGTTCTAAAAAAACC 9360
Db 9301 ATTAAATATTTGAGGTGTTTTCACCACTTATATCTTTCAAGTTCTAAAAAAACC 9360

Db 9301 ATTAAATATTTGAGGTGTTTTCACCACTTATATCTTTCAAGTTCTAAAAAAACC 9360
Qy 9361 GTATATATGTTTAACTCTAACAAAACCTTATATATGTTTCTCTAATACAGGAGTGT 9420
Db 9361 GTATATATGTTTAACTCTAACAAAACCTTATATATGTTTCTCTAATACAGGAGTGT 9420
Qy 9421 TCTTGAATGGAATTTGGGTCGCTATTAATCTTTTGGCTGCTGCTCAATCTTCTCAAGA 9480
Db 9421 TCTTGAATGGAATTTGGGTCGCTATTAATCTTTTGGCTGCTGCTCAATCTTCTCAAGA 9480
Qy 9481 TGAAGTGAAGAGGTCAATATCTCTCAAGAGAACAGAGAAATCTATATATATG 9540
Db 9481 TGAAGTGAAGAGGTCAATATCTCTCAAGAGAACAGAGAAATCTATATATATG 9540
Qy 9541 TTTGATGTTGACACACCCCTAGTTGATCTGACTGAAGCTGTAATCATTTCTAGTT 9600
Db 9541 TTTGATGTTGACACACCCCTAGTTGATCTGACTGAAGCTGTAATCATTTCTAGTT 9600
Qy 9601 ATCCCATTTATATTTCAATTAATAAGATATGTAATGTTTGTGAGCTGCAAGT 9660
Db 9601 ATCCCATTTATATTTCAATTAATAAGATATGTAATGTTTGTGAGCTGCAAGT 9660
Qy 9661 CATGACTCTACTTTTGAATGATTTGGCTCATTTGCTGCAATCAGTTCAATTAAT 9720
Db 9661 CATGACTCTACTTTTGAATGATTTGGCTCATTTGCTGCAATCAGTTCAATTAAT 9720
Qy 9721 GGTGTGTACACATGATGATGAGAGCGAGGTGTTTGAACACCTTTTCAACAAAAT 9780
Db 9721 GGTGTGTACACATGATGATGAGAGCGAGGTGTTTGAACACCTTTTCAACAAAAT 9780
Qy 9781 CTATATCTTCAACAATGAAACCTTGAATCCCTTGAAGAGCAACATCTCTT 9840
Db 9781 CTATATCTTCAACAATGAAACCTTGAATCCCTTGAAGAGCAACATCTCTT 9840
Qy 9841 GAATATGCTATGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9900
Db 9841 GAATATGCTATGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 9900
Qy 9901 CATGCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9960
Db 9901 CATGCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9960
Qy 9961 ACATCTCTGAGATTAATGTCAGTTTGTGATGATGATGATGATGATGATGATGATG 10020
Db 9961 ACATCTCTGAGATTAATGTCAGTTTGTGATGATGATGATGATGATGATGATGATG 10020
Qy 10021 GGGGGGGGGTGAATAGCGGTTTATTAACAATGATATATGTAATCTTAAATGTGGA 10080
Db 10021 GGGGGGGGGTGAATAGCGGTTTATTAACAATGATATATGTAATCTTAAATGTGGA 10080
Qy 10081 ATTAACTAGGTGATATTTTTCCAATTAAGGTCCTTTTATGACTCACATGTACCA 10140
Db 10081 ATTAACTAGGTGATATTTTTCCAATTAAGGTCCTTTTATGACTCACATGTACCA 10140
Qy 10141 TCAAGGATCAATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10200
Db 10141 TCAAGGATCAATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10200
Qy 10201 CCAACACACACACGACACACATTTCAACACCAATCAATGCTGACGAAGAGGAA 10260
Db 10201 CCAACACACACACGACACACATTTCAACACCAATCAATGCTGACGAAGAGGAA 10260
Qy 10261 GTCATACAAATCAAAAATGCTGCTAGGCGGAGAGAAATGAAATCATGAAAGAAATGA 10320
Db 10261 GTCATACAAATCAAAAATGCTGCTAGGCGGAGAGAAATGAAATCATGAAAGAAATGA 10320
Qy 10321 AAAACGCTGACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10380
Db 10321 AAAACGCTGACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10380
Qy 10381 CTGCGAGAAAAGTTCTATTAATACTGGGCTTGTGAGAGAGAAACGAGTCAAGAGTTGC 10440
Db 10381 CTGCGAGAAAAGTTCTATTAATACTGGGCTTGTGAGAGAGAAACGAGTCAAGAGTTGC 10440

QY	10441	CATCATCGGATCCAAACCACTAAGGTCATATCCTGGGTTTATCTCTGAAGATCAAAATCCG	1050
Db	10441	CATCATCGGATCCAAACCACTAAGGTCATATCCTGGGTTTATCTCTGAAGATCAAAATCCG	1050
QY	10501	AGCAAACTCCGAGTAATGCTTTTATTTAGGTTACGATTCAAAATAATGCCAAATCATGAG	1056
Db	10501	AGCAAACTCCGAGTAATGCTTTTATTTAGGTTACGATTCAAAATAATGCCAAATCATGAG	1056
QY	10561	TTATGACCAATTGAGCAGACAGCTAGAGATTTTATATCAAAAGCTCGAGCGGGTATCTTGA	1062
Db	10561	TTATGACCAATTGAGCAGACAGCTAGAGATTTTATATCAAAAGCTCGAGCGGGTATCTTGA	1062
QY	10621	AGTACCATTCATTAATGAGTCATCCACTTGCCTCAATAAATAGTTGCATAGATGACG	1068
Db	10621	AGTACCATTCATTAATGAGTCATCCACTTGCCTCAATAAATAGTTGCATAGATGACG	1068
QY	10661	GTCCATATGGGAGTAATGACATGAGAGGGCCATGTGTAAAGTTAACTGACGTGACAGA	1074
Db	10661	GTCCATATGGGAGTAATGACATGAGAGGGCCATGTGTAAAGTTAACTGACGTGACAGA	1074
QY	10741	GCCTGTCCGACCACTCGACAGAGTGTTTGAATGGGAGAAAGATATGCTCCACCAAC	1080
Db	10741	GCCTGTCCGACCACTCGACAGAGTGTTTGAATGGGAGAAAGATATGCTCCACCAAC	1080
QY	10801	ATCCCAAGTTGAAAACATTTCTTAGAGCCCTTATCCATCTATCAAAAGCACAATGATGAC	1086
Db	10801	ATCCCAAGTTGAAAACATTTCTTAGAGCCCTTATCCATCTATCAAAAGCACAATGATGAC	1086
QY	10861	TATCTGTATCAGACGACAAATCCATGTCCTCGATCATCGCTCTATCTATGCTCATTTGACATAC	1092
Db	10861	TATCTGTATCAGACGACAAATCCATGTCCTCGATCATCGCTCTATCTATGCTCATTTGACATAC	1092
QY	10921	TACCTGGCAAGGCGGATTCAGGCCCGACAGACGCTGGGCGGCGCG	1096
Db	10921	TACCTGGCAAGGCGGATTCAGGCCCGACAGACGCTGGGCGGCGCG	1096
RESULT 3			
AB024006			
LOCUS			
DEFINITION	AB024006	Hordeum vulgare naat-B and naat-A genes for nicotianamine aminotransferase, complete cds.	
ACCESSION	AB024006.1	GI:6469088	
VERSION			
KEYWORDS		nicotianamine aminotransferase.	
SOURCE		Hordeum vulgare subsp. vulgare	
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poidea; Triticaceae; Hordeum.	
REFERENCE		1 (bases)	
AUTHORS		Takahashi,M., Yamaguchi,H., Nakaniishi,H., Shidori,T., Nishizawa,N.K. and Mori,S.	
TITLE		Cloning two genes for nicotianamine aminotransferase, a critical enzyme in iron acquisition (Strategy II) in graminaceous plants	
JOURNAL		Plant Physiol. 121 (3), 947-956 (1999)	
REFERENCE		2 (bases 1 to 10966)	
AUTHORS		Mori,S., Takahashi,M. and Nakaniishi,H.	
TITLE		Direct Submision	
AUTHORS		Submitted (22-FEB-1999) Satoshi Mori, University of Tokyo, Dept. of Applied Biological Chemistry, Yayoi 1-1, Bunkyo-ku, Tokyo 113-8657, Japan (E-mail:samori@hongo.ecc.u-tokyo.ac.jp, Tel:81-3-3812-2111 (ex.5106), Fax:81-3-3812-0544)	
FEATURES		Location/Qualifiers	
SOURCE		1..10966	
		/organism="Hordeum vulgare subsp. vulgare"	
		/mol_type="genomic DNA"	
		/sub_features="1g1"	
		/db_xref="taxon:112509"	
		/clone_lib="lambdaFIXII"	

Query Match	100.0%;	Score 10966;	DB 8;	Length 10966;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 10966;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	CTCGATCCCATGCGATGTTATGATTAGTATCAACGAAAGAAAGATGCGATGTGCC	60	
Db	1	CTCGATCCCATGCGATGTTATGATTAGTATCAACGAAAGAAAGATGCGATGTGCC	60	
QY	61	CTGTGTGTCAATCCCTCACTGGCTTTGGGAGATGGCGATACCGGATTTGGTGAAGTGT	120	
Db	61	CTGTGTGTCAATCCCTCACTGGCTTTGGGAGATGGCGATACCGGATTTGGTGAAGTGT	120	
QY	121	TTAGCATGATGTCTGCGCGACCTGCGCAAGAAACTGGCGTGCACGCGATCTGCAGAGA	180	
Db	121	TTAGCATGATGTCTGCGCGACCTGCGCAAGAAACTGGCGTGCACGCGATCTGCAGAGA	180	
QY	181	GAGCGATGCAATGCTTTGTGTATGACCGGAGCTGAGTGGGTCTCACTAACTGAACCCCA	240	
Db	181	GAGCGATGCAATGCTTTGTGTATGACCGGAGCTGAGTGGGTCTCACTAACTGAACCCCA	240	
QY	241	GCATTGGGTAAGTTCGAGTTCGAGAGCCTCACTGCTTCCGCGTCCCGATTCGCTTATCT	300	
Db	241	GCATTGGGTAAGTTCGAGTTCGAGAGCCTCACTGCTTCCGCGTCCCGATTCGCTTATCT	300	
QY	301	TTCTCCCAATTTATTAAGAGGATAGATGATGCTGTGCTGGATTGGATAGTACGTGC	360	
Db	301	TTCTCCCAATTTATTAAGAGGATAGATGATGCTGTGCTGGATTGGATAGTACGTGC	360	
QY	361	ATTGAACCAAGAGGCGCCGAAATATTCGCGGAGATTAATGGTGCAGGCGCCCAACGCG	420	
Db	361	ATTGAACCAAGAGGCGCCGAAATATTCGCGGAGATTAATGGTGCAGGCGCCCAACGCG	420	

421 CACGCCCTCAGCTGCGAGCGGCGTCCAGAGCGTCCAGAGCGTCCGCGTGGTCT 480
421 CACGCCCTCAGCTGCGAGCGGCGTCCAGAGCGTCCAGAGCGTCCGCGTGGTCT 480
481 TCTTGTGCTCCGAGCGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540
481 TCTTGTGCTCCGAGCGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540
541 ACCGCTACTCAATCGTCCGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
541 ACCGCTACTCAATCGTCCGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
601 GTAGTCTAGTACT 660
601 GTAGTCTAGTACT 660
661 ACGCCAGAGCGAGCGAGTGGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 720
661 ACGCCAGAGCGAGCGAGTGGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 720
721 GAGCAACGCGAGCGAGTGGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 780
721 GAGCAACGCGAGCGAGTGGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 780
781 TGCCGACGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 840
781 TGCCGACGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 840
841 CAAAGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 900
841 CAAAGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 900
901 CAAAGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 960
901 CAAAGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 960
961 CGAGCATGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1020
961 CGAGCATGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1020
1021 AG 1080
1021 AG 1080
1081 GACGCGCGAG 1140
1081 GACGCGCGAG 1140
1141 GAAAG 1200
1141 GAAAG 1200
1201 CCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1260
1201 CCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1260
1261 CTGCTACCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1320
1261 CTGCTACCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1320
1321 GAAAG 1380
1321 GAAAG 1380
1381 TTGAG 1440
1381 TTGAG 1440
1441 GACCTTCTCCAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1500
1441 GACCTTCTCCAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1500
1501 GTCCAG 1560
1501 GTCCAG 1560

1501 GTCCAG 1560
1561 CCAAGCGAG 1620
1561 CCAAGCGAG 1620
1621 CCAAGCGAG 1680
1621 CCAAGCGAG 1680
1681 GCAATTTGAG 1740
1681 GCAATTTGAG 1740
1741 GCGCGAG 1800
1741 GCGCGAG 1800
1801 TTACTCTGAG 1860
1801 TTACTCTGAG 1860
1861 GTTCAAG 1920
1861 GTTCAAG 1920
1921 GGTGCGAG 1980
1921 GGTGCGAG 1980
1981 GGTGCGAG 2040
1981 GGTGCGAG 2040
2041 GGTGCGAG 2100
2041 GGTGCGAG 2100
2101 GGTGCGAG 2160
2101 GGTGCGAG 2160
2161 GGTGCGAG 2220
2161 GGTGCGAG 2220
2221 GGTGCGAG 2280
2221 GGTGCGAG 2280
2281 GGTGCGAG 2340
2281 GGTGCGAG 2340
2341 GGTGCGAG 2400
2341 GGTGCGAG 2400
2401 GGTGCGAG 2460
2401 GGTGCGAG 2460
2461 GGTGCGAG 2520
2461 GGTGCGAG 2520
2521 GGTGCGAG 2580
2521 GGTGCGAG 2580
2581 GGTGCGAG 2640
2581 GGTGCGAG 2640

Dh 2581 CAGCCATATTCCTTTGGGTCAGATGACAGTATGTTGGAACCTCTTTTAGAAAAATAGAT 2640
Qy 2641 ACATTTAGTATGATGATACCATTTTATTAAGAAATTCATATTTTGAATCCTGATGGT 2700
Dh 2641 ACATTTAGTATGATGATACCATTTTATTAAGAAATTCATATTTTGAATCCTGATGGT 2700
Qy 2701 ATTTGCTCTGATGATTCACAGATTTAATTGAGTTTATTTGATCTATCAAAATTTGTCAG 2760
Dh 2701 ATTTGCTCTGATGATTCACAGATTTAATTGAGTTTATTTGATCTATCAAAATTTGTCAG 2760
Qy 2761 GCAGCTCTGCTGATGATCTTGAAGAACCAAGGAAGTTTCTTTAAGGCGATATTTGGT 2820
Dh 2761 GCAGCTCTGCTGATGATCTTGAAGAACCAAGGAAGTTTCTTTAAGGCGATATTTGGT 2820
Qy 2821 CTGCTAAAGGAATCATCGAGATATGCTAACAAATTAAGAAAAAATAATACATTACA 2880
Dh 2821 CTGCTAAAGGAATCATCGAGATATGCTAACAAATTAAGAAAAAATAATACATTACA 2880
Qy 2881 TGTCTCTCAAGACCAAGGAAGATCAATGTTTGTCAATGAGCTATTTTGTGAATGAAA 2940
Dh 2881 TGTCTCTCAAGACCAAGGAAGATCAATGTTTGTCAATGAGCTATTTTGTGAATGAAA 2940
Qy 2941 AAATCTTAGGAGATGTCAGTATCATTAACCTATTTATATAGATTAATCTGGGACCGAA 3000
Dh 2941 AAATCTTAGGAGATGTCAGTATCATTAACCTATTTATATAGATTAATCTGGGACCGAA 3000
Qy 3001 ATGATCAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
Dh 3001 ATGATCAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
Qy 3061 TGAAGATGATGATTTTAAAGATTAAGACAGAGCTAAAGTTATGCTACACATCAATCAT 3120
Dh 3061 TGAAGATGATGATTTTAAAGATTAAGACAGAGCTAAAGTTATGCTACACATCAATCAT 3120
Qy 3121 CTGAGATATCCCTTTCATCTTTTGTATATCTGAGAAATGTTATGCTCAATCATATTT 3180
Dh 3121 CTGAGATATCCCTTTCATCTTTTGTATATCTGAGAAATGTTATGCTCAATCATATTT 3180
Qy 3181 ACACAAAATATCCATGATGATTTCTAGAAAGATGATTAATTTCTGTAACCATGATCTCC 3240
Dh 3181 ACACAAAATATCCATGATGATTTCTAGAAAGATGATTAATTTCTGTAACCATGATCTCC 3240
Qy 3241 GATATTAATCTTCATGATTAACAGGTTAACTGAACTTACATTTTGGAGAAATAGACGA 3300
Dh 3241 GATATTAATCTTCATGATTAACAGGTTAACTGAACTTACATTTTGGAGAAATAGACGA 3300
Qy 3301 TGACATTTGATTTTGTGCAAGCTCGCAAAAGAAATCAATTAATTTCCGAGATGACATA 3360
Dh 3301 TGACATTTGATTTTGTGCAAGCTCGCAAAAGAAATCAATTAATTTCCGAGATGACATA 3360
Qy 3361 GAATCCATTTGATTTTGTGCAAGCTCGCAAAAGAAATCAATTAATTTCCGAGATGACATA 3420
Dh 3361 GAATCCATTTGATTTTGTGCAAGCTCGCAAAAGAAATCAATTAATTTCCGAGATGACATA 3420
Qy 3421 CATTTAAATGATTTTCAATTTATGCTGAAATTTGTTAACTGATTAATTTCTGCTG 3480
Dh 3421 CATTTAAATGATTTTCAATTTATGCTGAAATTTGTTAACTGATTAATTTCTGCTG 3480
Qy 3481 AAGACTTTTAAACACAATATTTCTGTAAGATGATGATGATGATGATGATGATGATG 3540
Dh 3481 AAGACTTTTAAACACAATATTTCTGTAAGATGATGATGATGATGATGATGATGATG 3540
Qy 3541 TTGACCTGATGCTCTGCTCAATGTTGTGCTCTCAAAAAAATGACAGACATGTT 3600
Dh 3541 TTGACCTGATGCTCTGCTCAATGTTGTGCTCTCAAAAAAATGACAGACATGTT 3600
Qy 3601 TCAAAATATTTGTAATTAATTAATTTAGCCTTTAATTAATTTAGCCTTTAATTTAG 3660
Dh 3601 TCAAAATATTTGTAATTAATTAATTTAGCCTTTAATTAATTTAGCCTTTAATTTAG 3660
Qy 3661 TTTTGTAGTCAACTTATATTTTGTGACATTTCTCAAAAAAATAATTAATTTGTG 3720
Dh 3661 TTTTGTAGTCAACTTATATTTTGTGACATTTCTCAAAAAAATAATTAATTTGTG 3720

Qy 3721 CCTCAAAATGTAAGGAGTGTCTTGAATGCAAACTGGGTCGACATTAATTTTGTCTG 3780
Dh 3721 CCTCAAAATGTAAGGAGTGTCTTGAATGCAAACTGGGTCGACATTAATTTTGTCTG 3780
Qy 3781 TGTTCATCTCTCTTCAAGATGCTCGGAAGGATCAAAATCTTGTCAAAAGAA 3840
Dh 3781 TGTTCATCTCTCTTCAAGATGCTCGGAAGGATCAAAATCTTGTCAAAAGAA 3840
Qy 3841 GAAAGAAATTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 3900
Dh 3841 GAAAGAAATTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 3900
Qy 3901 CAGATCCCATGATTAATCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960
Dh 3901 CAGATCCCATGATTAATCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960
Qy 3961 CAACATCTCTTGAATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
Dh 3961 CAACATCTCTTGAATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
Qy 4021 GTTAGTAAATGATCAATGAGGCTGAGAGATCTTTGAAGTTTATTTTATTTTGGGGGG 4080
Dh 4021 GTTAGTAAATGATCAATGAGGCTGAGAGATCTTTGAAGTTTATTTTATTTTGGGGGG 4080
Qy 4081 GGGGGGGGGGGCTTGAATATTAATCTTTAAGTACAGTTCTCAAGTTATGTCMAAGCA 4140
Dh 4081 GGGGGGGGGGGCTTGAATATTAATCTTTAAGTACAGTTCTCAAGTTATGTCMAAGCA 4140
Qy 4141 CTTTGTAAATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 4200
Dh 4141 CTTTGTAAATTTGATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 4200
Qy 4201 AAGCAAAACCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4260
Dh 4201 AAGCAAAACCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4260
Qy 4261 GATCAGTTGATGATATCCCAATCAATGAGTAAATCATGTTGTTGCTACCACTTTT 4320
Dh 4261 GATCAGTTGATGATATCCCAATCAATGAGTAAATCATGTTGTTGCTACCACTTTT 4320
Qy 4321 CTACATTCCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4380
Dh 4321 CTACATTCCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4380
Qy 4381 CTCTGTGCAAAATTCGGAACCAAGATTAATCATCTCAAGCTGGAAGCAACCAAAATTTG 4440
Dh 4381 CTCTGTGCAAAATTCGGAACCAAGATTAATCATCTCAAGCTGGAAGCAACCAAAATTTG 4440
Qy 4441 AAGAGAGAAACCTTCAAAAGATCTTTGATGCTCCCAATGATGATGATGATGATGAT 4500
Dh 4441 AAGAGAGAAACCTTCAAAAGATCTTTGATGCTCCCAATGATGATGATGATGATGAT 4500
Qy 4501 TCGCAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
Dh 4501 TCGCAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
Qy 4561 CGATTAATGATTTTCCCAAAATCTTTAATTTGATGATGATGATGATGATGATGATGAT 4620
Dh 4561 CGATTAATGATTTTCCCAAAATCTTTAATTTGATGATGATGATGATGATGATGATGAT 4620
Qy 4621 ATTTTGAATTTGTAACAAATTTTAAACGGGTAATCTGTAACATTTTCAAAATTTG 4680
Dh 4621 ATTTTGAATTTGTAACAAATTTTAAACGGGTAATCTGTAACATTTTCAAAATTTG 4680
Qy 4681 GATCAAAATTTTAAACGAATTTTCTCAAAATTTTAAACGAATTTTAAACGAATTTT 4740
Dh 4681 GATCAAAATTTTAAACGAATTTTCTCAAAATTTTAAACGAATTTTAAACGAATTTT 4740
Qy 4741 GTTCAACAAATTTTGAATTTTAAATTAAGCAGAAATTTTGAATTTTAAATTAATTTT 4800
Dh 4741 GTTCAACAAATTTTGAATTTTAAATTAAGCAGAAATTTTGAATTTTAAATTAATTTT 4800

QY 4801 CGAATTTGGAACATTTTCTATTCTGACAAAAATTGAAATATACGAACGTAATTTGGA 4860
|
|
|
Db 4801 CGAATTTGGAACATTTTCTATTCTGACAAAAATTGAAATATACGAACGTAATTTGGA 4860
|
QY 4861 ATTAATTTTGGAAAAATGCGATTTTGTGAAATTTCTGAAATTTTGAATAAAACAAAAA 4920
|
|
|
Db 4861 ATTAATTTTGGAAAAATGCGATTTTGTGAAATTTCTGAAATTTTGAATAAAACAAAAA 4920
|
QY 4921 CTTTAAAGGTAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 4980
|
|
|
Db 4921 CTTTAAAGGTAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 4980
|
QY 4981 AAGAAATCCGAGAAAAAGCCAACTGGGAATGACATGGAATAAACCAGCGCTCGCGCCG 5040
|
|
|
Db 4981 AAGAAATCCGAGAAAAAGCCAACTGGGAATGACATGGAATAAACCAGCGCTCGCGCCG 5040
|
QY 5041 ACTGTGTAAGCTATTAAGTAGAGCCGCGCCAAAGCTCGTCTCATCATCATCTGTGCGA 5100
|
|
|
Db 5041 ACTGTGTAAGCTATTAAGTAGAGCCGCGCCAAAGCTCGTCTCATCATCATCTGTGCGA 5100
|
QY 5101 AACCCCGCAATTTGTTGCTGACATATGCGCGCAATAGAGCTTTTCAGAGAGCTCTGTCTCC 5160
|
|
|
Db 5101 AACCCCGCAATTTGTTGCTGACATATGCGCGCAATAGAGCTTTTCAGAGAGCTCTGTCTCC 5160
|
QY 5161 GGTATGAGGTCAATTGACACACCCCTCTGCACTTGAGGCGAGGCTATTAATCTTTTTC 5220
|
|
|
Db 5161 GGTATGAGGTCAATTGACACACCCCTCTGCACTTGAGGCGAGGCTATTAATCTTTTTC 5220
|
QY 5221 TTCTTTGCACTCAAGTTACTACGCGCAGTTAGTTTGTGAAAGCGAACCAACCGGTTTTGT 5280
|
|
|
Db 5221 TTCTTTGCACTCAAGTTACTACGCGCAGTTAGTTTGTGAAAGCGAACCAACCGGTTTTGT 5280
|
QY 5281 GAAAGTTCTAAGAACTCAACCACTTTTGGGAAGCTTCTAGAAAGCTATGATGTTTCTTT 5340
|
|
|
Db 5281 GAAAGTTCTAAGAACTCAACCACTTTTGGGAAGCTTCTAGAAAGCTATGATGTTTCTTT 5340
|
QY 5341 TGACATGTAATTAATTTGATTTTCTTTCAATGACAACTTTTCAATTCAT 5400
|
|
|
Db 5341 TGACATGTAATTAATTTGATTTTCTTTCAATGACAACTTTTCAATTCAT 5400
|
QY 5401 GATTTTGTGAAACCTGTGATTTTGTGAAATCGGTGATTTTCTTAATCGGTGTTT 5460
|
|
|
Db 5401 GATTTTGTGAAACCTGTGATTTTGTGAAATCGGTGATTTTCTTAATCGGTGTTT 5460
|
QY 5461 GAAAAAAAATGTCGATTTTCCGAATTAATGAACATTTGTTGCAAGATGAGATGATCT 5520
|
|
|
Db 5461 GAAAAAAAATGTCGATTTTCCGAATTAATGAACATTTGTTGCAAGATGAGATGATCT 5520
|
QY 5521 TTTCAATGAGCGAATTTTCTAAATATCAACATTTTCAATTCATAGCTTCC 5580
|
|
|
Db 5521 TTTCAATGAGCGAATTTTCTAAATATCAACATTTTCAATTCATAGCTTCC 5580
|
QY 5581 TTTTAAATGTCGTAATCTTATGCAATTTGTCGTAATCTTTTATTAATTAATGA 5640
|
|
|
Db 5581 TTTTAAATGTCGTAATCTTATGCAATTTGTCGTAATCTTTTATTAATTAATGA 5640
|
QY 5641 TTTTAAATGTCGTAATCTTATGCAATTTGTCGTAATCTTTTATTAATTAATGA 5700
|
|
|
Db 5641 TTTTAAATGTCGTAATCTTATGCAATTTGTCGTAATCTTTTATTAATTAATGA 5700
|
QY 5701 AACCATTTGACCTGCTGAAACAGGAGGCTTTCATATGATGAGGTGCTTAATACAGCG 5760
|
|
|
Db 5701 AACCATTTGACCTGCTGAAACAGGAGGCTTTCATATGATGAGGTGCTTAATACAGCG 5760
|
QY 5761 CCGCTGACTACTAAACGAAGGAATTTGCAATTTTACCAACCACTATGCTAATAAATG 5820
|
|
|
Db 5761 CCGCTGACTACTAAACGAAGGAATTTGCAATTTTACCAACCACTATGCTAATAAATG 5820
|
QY 5821 AATATTCGATTAATAAAGGGGAAAAAATCTAATCCCTGAAATCCCTCTCTCTTAAT 5880
|
|
|
Db 5821 AATATTCGATTAATAAAGGGGAAAAAATCTAATCCCTGAAATCCCTCTCTCTTAAT 5880
|
QY 5881 ATTTGTTGTTGGGAGAACTAATCTGAAGAATTAATCTAGTTCTCCGCAATAACAATA 5940

Db 5881 ATTTGTTGTTGGGAGAACTAATCTGAAGAATCTAATCTAGTTCTCCGCAATAACAATA 5940
|
|
|
QY 5941 TTATGATTTCCGGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6000
|
|
|
Db 5941 TTATGATTTCCGGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6000
|
QY 6001 ACCCAAGAAAGTCTGAGTTTGTGTTTGAAGAACGGAAGATGATGATGCTTCTGA 6060
|
|
|
Db 6001 ACCCAAGAAAGTCTGAGTTTGTGTTTGAAGAACGGAAGATGATGATGCTTCTGA 6060
|
QY 6061 ATTAGTTCATCTTTTCCCAAGATTAAGAAGATGAATTTGATTAATGATGCTG 6120
|
|
|
Db 6061 ATTAGTTCATCTTTTCCCAAGATTAAGAAGATGAATTTGATTAATGATGCTG 6120
|
QY 6121 GAGATATCATTTCTGATTAAGTACAAATTTGTTTCCGGGACAGCCAAACCCGCACTGG 6180
|
|
|
Db 6121 GAGATATCATTTCTGATTAAGTACAAATTTGTTTCCGGGACAGCCAAACCCGCACTGG 6180
|
QY 6181 AGCCGGAATTGGAATTAAGTGGGTGAGTGCAGAAAGCATGTTGATGCTTCTCAAGAG 6240
|
|
|
Db 6181 AGCCGGAATTGGAATTAAGTGGGTGAGTGCAGAAAGCATGTTGATGCTTCTCAAGAG 6240
|
QY 6241 TGTAGCAGTATGCTGCTCTGTTGCTGAGCTGCAATTAAGATACATTAATAACAA 6300
|
|
|
Db 6241 TGTAGCAGTATGCTGCTCTGTTGCTGAGCTGCAATTAAGATACATTAATAACAA 6300
|
QY 6301 GACGATCAGCTGAGCGGTGCTGATGCTGCTTCTTCTGCTGCGCCCGGAGCCCGG 6360
|
|
|
Db 6301 GACGATCAGCTGAGCGGTGCTGATGCTGCTTCTTCTGCTGCGCCCGGAGCCCGG 6360
|
QY 6361 TTGATGTCGCGAGCGAGTGGCGGACGGAACCGGCTATTAACAGGCGCAACGCAACGCG 6420
|
|
|
Db 6361 TTGATGTCGCGAGCGAGTGGCGGACGGAACCGGCTATTAACAGGCGCAACGCAACGCG 6420
|
QY 6421 CGCGTTCTCAATCATCATCATCTTCTGATGTTGATTTGATCTAGTATTCCTCTG 6480
|
|
|
Db 6421 CGCGTTCTCAATCATCATCATCTTCTGATGTTGATTTGATCTAGTATTCCTCTG 6480
|
QY 6481 CCACATCTGATGATCT 6540
|
|
|
Db 6481 CCACATCTGATGATCT 6540
|
QY 6541 CGCGGAGGCG 6600
|
|
|
Db 6541 CGCGGAGGCG 6600
|
QY 6601 CGCGGAGGCG 6660
|
|
|
Db 6601 CGCGGAGGCG 6660
|
QY 6661 GAGCGGATCTCTGCGGACGAGCGGGGCGAAGAACGATCCGCGGCGATCAAGAT 6720
|
|
|
Db 6661 GAGCGGATCTCTGCGGACGAGCGGGGCGAAGAACGATCCGCGGCGATCAAGAT 6720
|
QY 6721 CAGGCGAGGCTGAGAGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6780
|
|
|
Db 6721 CAGGCGAGGCTGAGAGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6780
|
QY 6781 GTCCGTTTCCCGGCTTTCGCGACGCGCGCTGAGAGCGGCGCGCGCGCGCGCGCT 6840
|
|
|
Db 6781 GTCCGTTTCCCGGCTTTCGCGACGCGCGCTGAGAGCGGCGCGCGCGCGCGCGCT 6840
|
QY 6841 GCGGACCGGCGGATTAATGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6900
|
|
|
Db 6841 GCGGACCGGCGGATTAATGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6900
|
QY 6901 ACATTTACAGCTTACAGGTAATGATGATGATGATGATGATGATGATGATGATGATG 6960
|
|
|
Db 6901 ACATTTACAGCTTACAGGTAATGATGATGATGATGATGATGATGATGATGATGATG 6960
|
QY 6961 CCGCGGCTGTTCTTCCCGGCTGCGTTCAAAATTTTAACTTTTAACTTTTAACTTTTAA 7020
|
|
|

Dh 6961 CCGCGCTGTTCTTCCCGGTGCGCTCAAAATTTTAACTTCTAATAGTACTTATATAA 7020
Qy 7021 ACAAACGCGCGCTAGAGAGACACTTGTCAAGGGCGCGCTTCAAGCTATGCGCGGAC 7080
Dh 7021 ACAAACGCGCGCTAGAGAGACACTTGTCAAGGGCGCGCTTCAAGCTATGCGCGGAC 7080
Qy 7081 GAGCTTCTTCTACCGCGCGGGAATCTGAGCGATCGAAGTCAATATCCCGGTCTGGCC 7140
Dh 7081 GAGCTTCTTCTACCGCGCGGGAATCTGAGCGATCGAAGTCAATATCCCGGTCTGGCC 7140
Qy 7141 CAGAGTCCGCGCGCAATATCTGCTTCCCGCGCAAGCTATCCAAATTAAGAGCGCGA 7200
Dh 7141 CAGAGTCCGCGCGCAATATCTGCTTCCCGCGCAAGCTATCCAAATTAAGAGCGCGA 7200
Qy 7201 GCGGCAATTCACAAAGCTGAGAGTCCGGCACTTCGACCTCATCCCGCAAGGGGTGGAG 7260
Dh 7201 GCGGCAATTCACAAAGCTGAGAGTCCGGCACTTCGACCTCATCCCGCAAGGGGTGGAG 7260
Qy 7261 ATCGACATCGACTGCTGGAATCCATGCGCGCAAGAACACACCGCGATGCTCATATA 7320
Dh 7261 ATCGACATCGACTGCTGGAATCCATGCGCGCAAGAACACACCGCGATGCTCATATA 7320
Qy 7321 AACCCAAACATCCGTCGCGGCAAGGTTTACTCTTCAACCATCTGCGCAAGGTTTGAT 7380
Dh 7321 AACCCAAACATCCGTCGCGGCAAGGTTTACTCTTCAACCATCTGCGCAAGGTTTGAT 7380
Qy 7381 CCATGACCTCTGCTGCTGATGATCGACCGGCTGTTGAACATAGTATGATGATGCT 7440
Dh 7381 CCATGACCTCTGCTGCTGATGATCGACCGGCTGTTGAACATAGTATGATGATGCT 7440
Qy 7441 TTGCTAATCGTGTCTGATGATGATGCTGTTGTTATCAAGTGGCGAGGTGGCAAGAAC 7500
Dh 7441 TTGCTAATCGTGTCTGATGATGATGCTGTTGTTATCAAGTGGCGAGGTGGCAAGAAC 7500
Qy 7501 TCGGAATATGAGTATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 7560
Dh 7501 TCGGAATATGAGTATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 7560
Qy 7561 TTATCCCGATGGGCTCTTGGGCACTTGGCGCGGCTTGTCCATGATGATGATGATGAT 7620
Dh 7561 TTATCCCGATGGGCTCTTGGGCACTTGGCGCGGCTTGTCCATGATGATGATGATGAT 7620
Qy 7621 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7680
Dh 7621 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7680
Qy 7681 TTTTGAAGAAATCTAAGTACTTAACTCCCTAATCTTCTCAATGCTAATGCTAATGCT 7740
Dh 7681 TTTTGAAGAAATCTAAGTACTTAACTCCCTAATCTTCTCAATGCTAATGCTAATGCT 7740
Qy 7741 GATTAGTATTTTGTCTAATTTTGTACTGCTTGTGTTATTCAGATCTCTAATGCTAAT 7800
Dh 7741 GATTAGTATTTTGTCTAATTTTGTACTGCTTGTGTTATTCAGATCTCTAATGCTAAT 7800
Qy 7801 GAATTAATCTTAATGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 7860
Dh 7801 GAATTAATCTTAATGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 7860
Qy 7861 GAAATTAATCTTAATGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 7920
Dh 7861 GAAATTAATCTTAATGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 7920
Qy 7921 AGAGTTTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 7980
Dh 7921 AGAGTTTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 7980
Qy 7981 TTTAGTCAATCCCTTGTGTTGATTCATTAATAATTAATAATAATAATAATAATAATA 8040
Dh 7981 TTTAGTCAATCCCTTGTGTTGATTCATTAATAATTAATAATAATAATAATAATAATA 8040
Qy 8041 AACATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 8100
Dh 8041 AACATTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 8100

Qy 8101 TCTTTGATCAAGTATGCAATATATTTAAACTTGTGATGATGATGATGATGATGATGAT 8160
Dh 8101 TCTTTGATCAAGTATGCAATATATTTAAACTTGTGATGATGATGATGATGATGATGAT 8160
Qy 8161 GTTAAGATTTCTTAACTTCAACACCTTGTGCAATCTTCAATTTGATATCTTGA 8220
Dh 8161 GTTAAGATTTCTTAACTTCAACACCTTGTGCAATCTTCAATTTGATATCTTGA 8220
Qy 8221 GATTATGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8280
Dh 8221 GATTATGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8280
Qy 8281 TTAGAGAGCTCTTCTAAATTTCTTGAAGACAAAGAGATTTCTTAAAGAGATTAAT 8340
Dh 8281 TTAGAGAGCTCTTCTAAATTTCTTGAAGACAAAGAGATTTCTTAAAGAGATTAAT 8340
Qy 8341 TGGTCTAATTAAGAAATCATCAGAGATGATGATGATGATGATGATGATGATGATGAT 8400
Dh 8341 TGGTCTAATTAAGAAATCATCAGAGATGATGATGATGATGATGATGATGATGATGAT 8400
Qy 8401 TACGTGCTCAAGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 8460
Dh 8401 TACGTGCTCAAGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 8460
Qy 8461 CTTTTAAGGTTAATCTGGATCTCAGTGCATCAACAAACATCAATCAAAATATTAAT 8520
Dh 8461 CTTTTAAGGTTAATCTGGATCTCAGTGCATCAACAAACATCAATCAAAATATTAAT 8520
Qy 8521 TATGTTTGTGATGATCTTTTGAAGATGATGATGATGATGATGATGATGATGATGAT 8580
Dh 8521 TATGTTTGTGATGATCTTTTGAAGATGATGATGATGATGATGATGATGATGATGAT 8580
Qy 8581 AAATTAATTAAGACATATCTGATATCATTTGATGATGATGATGATGATGATGATGAT 8640
Dh 8581 AAATTAATTAAGACATATCTGATATCATTTGATGATGATGATGATGATGATGATGAT 8640
Qy 8641 ATTGATACACTCAATCAATATTAACAAAGTTTCTCCATTTTATGATGATGATGAT 8700
Dh 8641 ATTGATACACTCAATCAATATTAACAAAGTTTCTCCATTTTATGATGATGATGAT 8700
Qy 8701 CTTTTATTAACATGATATTAATCTTCATGATTAACAGTCAACATTAATCTTTT 8760
Dh 8701 CTTTTATTAACATGATATTAATCTTCATGATTAACAGTCAACATTAATCTTTT 8760
Qy 8761 GAGAGATCATGAGACATTAATTTTGTGCAAGTGTGCAAGAGAGAGATCTGAT 8820
Dh 8761 GAGAGATCATGAGACATTAATTTTGTGCAAGTGTGCAAGAGAGAGATCTGAT 8820
Qy 8821 TTTATGTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8880
Dh 8821 TTTATGTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8880
Qy 8881 ATCTTCTGTTATTAATCAACAAATCTTCTCCATCTCTAAATTAATGCTTAACTTAAG 8940
Dh 8881 ATCTTCTGTTATTAATCAACAAATCTTCTCCATCTCTAAATTAATGCTTAACTTAAG 8940
Qy 8941 ACAATTTTATCTAAGTCTAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 9000
Dh 8941 ACAATTTTATCTAAGTCTAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 9000
Qy 9001 TATATATGTTGAGAAATTAAGTATGTTGATGATGATGATGATGATGATGATGATGAT 9060
Dh 9001 TATATATGTTGAGAAATTAAGTATGTTGATGATGATGATGATGATGATGATGATGAT 9060
Qy 9061 CATGTTCTTATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 9120
Dh 9061 CATGTTCTTATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 9120
Qy 9121 ACATTAATGATTAACATATCATTTGTGAGAAATGATTAACATTAATGATTAATCTTCT 9180
Dh 9121 ACATTAATGATTAACATATCATTTGTGAGAAATGATTAACATTAATGATTAATCTTCT 9180

Qy 9181 TTTTATGCAAGAGTAGCAGGAGTACTTTCTCTGACCTATGAGTCTTTTTC 9240
 Db 9181 TTTTATGCAAGAGTAGCAGGAGTACTTTCTCTGACCTATGAGTCTTTTTC 9240
 Qy 9241 TCGTTACATGTTTGTGCTTCTACCAAAAATATATATACCAAGACATGTTCCAAATGAT 9300
 Db 9241 TCGTTACATGTTTGTGCTTCTACCAAAAATATATATACCAAGACATGTTCCAAATGAT 9300
 Qy 9301 ATTATATATTTTGTGAGTGTCTTCAACCACTATATCTTTCATAGTTCTTAAATACC 9360
 Db 9301 ATTATATATTTTGTGAGTGTCTTCAACCACTATATCTTTCATAGTTCTTAAATACC 9360
 Qy 9361 GTATATATGTTTAACTCTACCAAAAATATATATGTTTCTCTCTAATACAGGAGTGT 9420
 Db 9361 GTATATATGTTTAACTCTACCAAAAATATATATGTTTCTCTCTAATACAGGAGTGT 9420
 Qy 9421 TCTTGGAAATGGAATTTGGGTCCGTATTAATCTTTTGGCTGCGCTTCATCTCTTCAAGA 9480
 Db 9421 TCTTGGAAATGGAATTTGGGTCCGTATTAATCTTTTGGCTGCGCTTCATCTCTTCAAGA 9480
 Qy 9481 TGGACTCGAAAGGCTCAATCTCTGCAAGAAAGAAAGAAAGAAATTTCTATATATG 9540
 Db 9481 TGGACTCGAAAGGCTCAATCTCTGCAAGAAAGAAAGAAAGAAATTTCTATATATG 9540
 Qy 9541 TTGTTAGTTTGTACACACCCCTAGTTGTACATCTGACCTGATGTAATCTATTTAGTT 9600
 Db 9541 TTGTTAGTTTGTACACACCCCTAGTTGTACATCTGACCTGATGTAATCTATTTAGTT 9600
 Qy 9601 ATCCCATTTTATATTTCAATTAATAAATTTGTAATGTTCTGTGTAGCTGTCCAGT 9660
 Db 9601 ATCCCATTTTATATTTCAATTAATAAATTTGTAATGTTCTGTGTAGCTGTCCAGT 9660
 Qy 9661 CATGTACTCTATCTTTTGTATGATTTTGGCTCTATGCTTGCATCACTTCAATTAAT 9720
 Db 9661 CATGTACTCTATCTTTTGTATGATTTTGGCTCTATGCTTGCATCACTTCAATTAAT 9720
 Qy 9721 GGTGTGTGTACCAATGATGATGTAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9780
 Db 9721 GGTGTGTGTACCAATGATGATGTAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9780
 Qy 9781 CTATATCTTTTCAACCAATGAAACCTTGAATTCCTTTGAGTGAAGTCAATCTCTT 9840
 Db 9781 CTATATCTTTTCAACCAATGAAACCTTGAATTCCTTTGAGTGAAGTCAATCTCTT 9840
 Qy 9841 GAATATGCTATGTTTCCATGCTGTGATGAAACATGATGAAATGAAATGAAATGATC 9900
 Db 9841 GAATATGCTATGTTTCCATGCTGTGATGAAACATGATGAAATGAAATGAAATGATC 9900
 Qy 9901 CATGTCAAAAGTTTATATGTTTAAATTCATTTATGAGAACTTGTATATCTTACGAC 9960
 Db 9901 CATGTCAAAAGTTTATATGTTTAAATTCATTTATGAGAACTTGTATATCTTACGAC 9960
 Qy 9961 ACATTTCTGTGAAGTATGTCAGTTTGTGTAAGGAGCCTATATTTTCTATTTGGG 10020
 Db 9961 ACATTTCTGTGAAGTATGTCAGTTTGTGTAAGGAGCCTATATTTTCTATTTGGG 10020
 Qy 10021 GGGGGGGGGTGAATGAGGGGTTTATACCAATTTGTAATTTGAGAAATCTTAAATG 10080
 Db 10021 GGGGGGGGGTGAATGAGGGGTTTATACCAATTTGTAATTTGAGAAATCTTAAATG 10080
 Qy 10081 ATTAACTAGGTGAATATTTTCCATTAAGGGGCTTTTATGAGCTCACAATGATCA 10140
 Db 10081 ATTAACTAGGTGAATATTTTCCATTAAGGGGCTTTTATGAGCTCACAATGATCA 10140
 Qy 10141 TCAAGGATATCAATATATATATGATGACAAATCTGATCTATATCTGATGATGACG 10200
 Db 10141 TCAAGGATATCAATATATATATGATGACAAATCTGATCTATATCTGATGATGACG 10200
 Qy 10201 CCAACACACACACACACACACACATTCACACACAAATCAATGCTGACCAAGGCGAA 10260
 Db 10201 CCAACACACACACACACACACACATTCACACACAAATCAATGCTGACCAAGGCGAA 10260
 Qy 10261 GTCATATCAAGATCAAAACTATGCTTGGCGGAGGAGAAATGAAAAATGAAAGAAATGA 10320

Db 10261 GTCATATCAAGATCAAAACTATGCTTACGCGGAGGAGAAATGAAAAATGAAATGA 10320
 Qy 10321 AAAACCGTACTGACCAATATCTGACATGACGACCAAACTCTGTGACCAACAAAA 10380
 Db 10321 AAAACCGTACTGACCAATATCTGACATGACGACCAAACTCTGTGACCAACAAAA 10380
 Qy 10381 CTGGGAGAAAAAGTCTATTAATCTGGGCGCTTGTGAGAAAGGAAAGAGGTGCAAGTTGC 10440
 Db 10381 CTGGGAGAAAAAGTCTATTAATCTGGGCGCTTGTGAGAAAGGAAAGAGGTGCAAGTTGC 10440
 Qy 10441 CATCATGCAATCCACCACTAGGTCAATCTGAGTTTTCATCTGAAATCAATCCG 10500
 Db 10441 CATCATGCAATCCACCACTAGGTCAATCTGAGTTTTCATCTGAAATCAATCCG 10500
 Qy 10501 AGCAAACTCCGAGTAATGCTTTTATGAGGTAAGATTCAAATAATGCAATCATGAG 10560
 Db 10501 AGCAAACTCCGAGTAATGCTTTTATGAGGTAAGATTCAAATAATGCAATCATGAG 10560
 Qy 10561 TTATGACCAATTTAGACGAGCTAGATTTTATGCAAGCTGAGACGGGTACTTGA 10620
 Db 10561 TTATGACCAATTTAGACGAGCTAGATTTTATGCAAGCTGAGACGGGTACTTGA 10620
 Qy 10621 AGTACCAATCCAAATTTGAATGATCCCACTTGCCTCAATACAAATGTTGACATGACG 10680
 Db 10621 AGTACCAATCCAAATTTGAATGATCCCACTTGCCTCAATACAAATGTTGACATGACG 10680
 Qy 10681 GTCCATATGCGAGTAATGACATGACGCGCATGTGTAGGTTTAACTGACATGACAA 10740
 Db 10681 GTCCATATGCGAGTAATGACATGACGCGCATGTGTAGGTTTAACTGACATGACAA 10740
 Qy 10741 GCTGTGCGCACTGACGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10800
 Db 10741 GCTGTGCGCACTGACGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10800
 Qy 10801 ATCCCAAGTTGAAACATTTAGAGCCCTTACCACTCAACAAAGGACAAATGATGAC 10860
 Db 10801 ATCCCAAGTTGAAACATTTAGAGCCCTTACCACTCAACAAAGGACAAATGATGAC 10860
 Qy 10861 TATCTGTATGACAGCAATTCATGTCCTGCTATCTTGTCTATGATGATGAC 10920
 Db 10861 TATCTGTATGACAGCAATTCATGTCCTGCTATCTTGTCTATGATGATGAC 10920
 Qy 10921 TACCTGGCAAAAGGCGGATTCAGCCCGACAGCAGCTGGGCGCGCC 10966
 Db 10921 TACCTGGCAAAAGGCGGATTCAGCCCGACAGCAGCTGGGCGCGCC 10966

RESULT 4
 AB005788
 LOCUS
 DEFINITION
 AB005788 1895 bp mRNA linear PLN 27-NOV-1999
 Hordeum vulgare mRNA for nicotianamine aminotransferase B, complete
 cds.
 ACCESSION
 AB005788
 VERSION
 AB005788.1 GI:6469086
 KEYWORDS
 nicotianamine aminotransferase B.
 SOURCE
 Hordeum vulgare subsp. vulgare
 ORGANISM
 Hordeum vulgare subsp. vulgare
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Pooideae; Triticaceae; Hordeum.
 REFERENCE
 1 (bases 1 to 1895)
 AUTHORS
 Takahashi, M., Yamaguchi, H., Nakanishi, H., Shioiri, T.,
 Nishizawa, N.K. and Mori, S.
 TITLE
 Cloning two genes for nicotianamine aminotransferase, a critical
 enzyme in iron acquisition (Strategy II) in Gramineae plants
 JOURNAL
 Plant Physiol. 121 (3), 947-956 (1999)
 MEDLINE
 20027437
 PUBMED
 10557244
 REFERENCES
 2 (bases 1 to 1895)
 AUTHORS
 Mori, S.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (08-JUL-1997) Satoshi Mori, The University of Tokyo,

Plant Molecular Physiology; Yayoi 1-1-1, Bunkyo-ku, Tokyo 113,
Japan (E-mail: aa1078@hongo.ecc.u-tokyo.ac.jp,
Tel:03-3812-2111 (ex. 5106), Fax:03-3812-0544)
Location/Qualifiers

FEATURES

Source

1..1895
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultiivar="Ehimedaka No.1"
/db_xref="taxon:112509"
/class type="root"
1..1895
/gene="naat-b"
76..1731
/gene="naat-b"
/note="NAAT-B"
/product="nicotianamine aminotransferase B"
/protein_id="BA87053.1"
/db_xref="GI:6469087"
/translation="MATVROSQDVAAAGLAVALAAANGKNGHVAALVNGKNGHVD
ADANGKNGHVAADANGKNGHAEATANGHGEATANGKNGHESNGHAEADANG
SNEHEDSANGSNGHAAAGAAAEBAEVENPFGAKDQVLAAGGANSIRAIYKISA
SVQKGRPRVPLAHDPSPVPAFRIVEADAVAAVTRGQENCTPAGTGLPARSA
VAEHLQGVPMASADVFLTAGSTQALEVILPVLAQTAGANILPFRPGVNYBARSA
FNRLVHFPLIPDKGWEIDLSLESIADKNTTAMVFINPNSGVSYSYDHSKVAE
VAKRIGLIVLADEVYKLVLSAPFIPMGVFGHTPVLSIGLSKSVYVGMRLKVAE
VYDPRKILQETKISTSTNYLVNSTDPAFTQALPQILENTKEDPFKILIGLKES
EICYKOIKENKTYTICPHKPGSMFVWVKMLNLIIEEDDDIDPCKLAKESVILCG
SVLGMANWVRITFCVPSLSLDGLGRKSKFCQRKRNKSSDDC"
1895
/gene="naat-b"
/note="15 A nucleotides"

polyA_site

ORIGIN

Query Match 8.9%; Score 971.8; DB 8; Length 1895;
Best Local Similarity 81.7%; Pred. No. 3.8e-115;
Matches 1287; Conservative 0; Mismatches 7; Indels 281; Gaps 3;

Qy 579 CGCTACTAGTAGTATTCCTGGTGTAGTCTAGTAGTA--CTCCCTCTCTCTCTCTCTCC 636
Db 3 CGCTACTAGTAGTATTCCTGGTGTAGTCTAGTAGTA--CTCTCTCTCTCTCTCTCTCC 62
Qy 637 TACCCGTTTCTCTCATGCGCAACCGTAACCGCAGAGCGAGTGCCTGCGAAACGGCTTGC 696
Db 63 TACCCGTTTCTCTCATGCGCAACCGTAACCGCAGAGCGAGTGCCTGCGAAACGGCTTGC 122
Qy 697 CGTGGCCGACGCGCGCGAAACCGCAAGAGCAACGGCAATGGCTGCTCCCGCGAAACGG 756
Db 123 CGTGGCCGACGCGCGCGAAACCGCAAGAGCAACGGCAATGGCTGCTCCCGCGAAACGG 182
Qy 757 CAAGAGCAACCGCATGGCTGTATGCGACGCGCAACGGCAACGGCAACGGCAATGGCT 816
Db 183 CAAGAGCAACCGCATGGCTGTATGCGACGCGCAACGGCAACGGCAACGGCAATGGCT 242
Qy 817 GCGTGCAGACGCGCAACCGCAAGAGCAACGGCAATGGCTGCGAAACGGCAACGG 876
Db 243 GCGTGCAGACGCGCAACCGCAAGAGCAACGGCAATGGCTGCGAAACGGCAACGG 302
Qy 877 CGAGGCAACGCGCAACCGCAAGAGCAACGGCAATGGCTGCGAAACGGCAACGG 936
Db 303 CGAGGCAACGCGCAACCGCAAGAGCAACGGCAATGGCTGCGAAACGGCAACGG 362
Qy 937 CGCGGACGCGCAACCGCAAGAGCAACGGCAATGGCTGCGAAACGGCAACGG 996
Db 363 CGCGGACGCGCAACCGCAAGAGCAACGGCAATGGCTGCGAAACGGCAACGG 422
Qy 997 CAACGGGCAATGGCTGCGCGCGCGCGCAAGAGAGAGAGCGGTGAGTGAATTTCCGCGG 1056
Db 423 CAACGGGCAATGGCTGCGCGCGCGCGCAAGAGAGAGAGCGGTGAGTGAATTTCCGCGG 482
Qy 1057 TGCCCAAGAGAGCGGTGCGCGCGCGCGCAAGAGAGAGAGCGGTGAGTGAATTTCCGCGG 1116
Db 483 TGCCCAAGAGAGCGGTGCGCGCGCGCGCAAGAGAGAGAGCGGTGAGTGAATTTCCGCGG 542

Qy 1117 CAAGATCAGCGCGAGAGGTGCAAGAGAGAGGCGCGCGCGCGCTGCTGCGCGCGCAACGG 1176
Db 543 CAAGATCAGCGCGAGAGGTGCAAGAGAGAGGCGCGCGCGCGCTGCTGCGCGCGCAACGG 602
Qy 1177 GAACCCGCTCGTGTTCCTGCGCTTCCGCAACCGCGCTGAGAGCGAGAGACCGCTGCGCGC 1236
Db 603 GAACCCGCTCGTGTTCCTGCGCTTCCGCAACCGCGCTGAGAGCGAGAGACCGCTGCGCGC 662
Qy 1237 CGCGCTGCGCAACCGGCAATGCTAACTGCTACCCCGCGCGGTGCGGCTTCCCGCGCGACG 1296
Db 663 CGCGCTGCGCAACCGGCAATGCTAACTGCTACCCCGCGCGGTGCGGCTTCCCGCGCGACG 722
Qy 1297 AAGTAAACAACAACAACAACAAGAACAAATTTCTTTTGGCTGTGCTGCGCGCGCA 1356
Db 723 A-----
Qy 1357 ATCCATGATGCGCATGTGCGCGCTTTCACGTGTCGTCGTCGTCACCGCTTCTTCTCT 1416
Db 724 -----
Qy 1417 CTTCCCTAGCGCCATGAGAAATCTGACCTTCTCCACCTTATACAAACAACAAAAA 1476
Db 724 -----
Qy 1477 ACAACGCGCGGTGCGAGAGCACTGTGCGAGGCGTGCCTGATCATGCTATCGGCGCGACA 1536
Db 724 -----AAGCGGTGCGAGAGCACTGTGCGAGGCGTGCCTGATCATGCTATCGGCGCGACA 779
Qy 1537 GGTCTTCTCTACCGCGCGCGGAGCCCAAGCGATGAGTGTATTAATCCGCTGCTGCCCA 1596
Db 780 GGTCTTCTCTACCGCGCGCGGAGCCCAAGCGATGAGTGTATTAATCCGCTGCTGCCCA 839
Qy 1597 GACCGCGCGCGCAACATCTGCTCCCGCGCGAGCTACCCAACTAGAGCGCGCGCG 1656
Db 840 GACCGCGCGCGCAACATCTGCTCCCGCGCGAGCTACCCAACTAGAGCGCGCGCGCG 899
Qy 1657 CGGTTTCAACAGGCTGAGAGTCCGCGATTTTCACTCATCCCGCAACAGGCGTGGAGAT 1716
Db 900 CGGTTTCAACAGGCTGAGAGTCCGCGATTTTCACTCATCCCGCAACAGGCGTGGAGAT 959
Qy 1717 CGACATGACTCGCTGGAATTCATCGCCGCAAGAACACCGCGCATGCTATCATTA 1776
Db 960 CGACATGACTCGCTGGAATTCATCGCCGCAAGAACACCGCGCATGCTATCATTA 1019
Qy 1777 CCGCAACAACCGGTGAGGAGAGGTTTACTCTTACGACCAATCTGCAAGGTTTCAATCC 1836
Db 1020 CCGCAACAACCGGTGAGGAGAGGTTTACTCTTACGACCAATCTGCAAGGTTTCAATCC 1066
Qy 1837 TTTGCTTGTGATATGATTCAGTTCAAGTCACTGCTGAATTTCTTTTTCGCAATCCG 1896
Db 1067 -----
Qy 1897 ATATCTGATCTGATTTGCTCAATTAAGTCCGCGAGGTGCGAAAGGCTCGGAATATTGGT 1956
Db 1067 -----AGGTCCGCGAGGTGCGAAAGGCTCGGAATATTGGT 1103
Qy 1957 GATTGCTGACGAGATTAACGCAAGCTGATTCGCGGACAGGCGCGCTTCAATCCCAATGGG 2016
Db 1104 GATTGCTGACGAGATTAACGCAAGCTGATTCGCGGACAGGCGCGCTTCAATCCCAATGGG 1163
Qy 2017 AGTGTGTGGCAATCAACCTGTGCTGTCAATAGGATCTTGTCCAAATCATGATAGT 2076
Db 1164 AGTGTGTGGCAATCAACCTGTGCTGTCAATAGGATCTTGTCCAAATCATGATAGT 1223
Qy 2077 GCTTGAATGGCGGCTTGTGATGGGTACCGGTGTACGACCCCAAGAAATCTTAACAGAAAC 2136
Db 1224 GCTTGAATGGCGGCTTGTGATGGGTACCGGTGTACGACCCCAAGAAATCTTAACAGAAAC 1283
Qy 2137 TAAGTACTTAATC 2151
Db 1284 TAAGTACTTAATC 1298

RESULT 5			
LOCUS	D88273		
DEFINITION	D88273	1660 bp	mRNA
ACCESSION	D88273		1linear
VERSION	D88273		PM 02-DEC-1996
KEYWORDS			
SOURCE	D88273.2	GI:6498121	
ORGANISM	neat-A; nicotianamine aminotransferase A.		
	Hordeum vulgare subsp. vulgare		
	Hordeum vulgare subsp. vulgare		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
	Poodeae, Triticeae; Hordeum.		
REFERENCE	1 (bases)		
AUTHORS	Takehashi, M., Yamaguchi, H., Nakanishi, H., Shiota, T.,		
	Nishizawa, N.K. and Mori, S.		
TITLE	Cloning two genes for nicotianamine aminotransferase, a critical		
	enzyme in iron acquisition (Strategy II) in graminaceous plants		
JOURNAL	Plant Physiol.	121 (3),	947-956 (1993)
MEDLINE	20027437		
PUBMED	10557244		
REFERENCE	2 (bases 1 to 1660)		
AUTHORS	Mori, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-OCT-1996)	Satoshi Mori,	The University of Tokyo,

COMMENT On Dec 1, 1999 this sequence version replaced gl:6469084 Sequence updated (27-Nov-1999).

FEATURES

SOURCE

gene	1. .1660
	/gene="naat-A"
CDS	62. .1447

[illegible]

ORIGIN

Query Match	6.8%;	Score 744.6;	DB 8;	Length 1660;
Best Local Similarity	80.6%;	Pred. No. 5.6e-86;		
Matches 1005; Conservative	0;	Mismatches 4;	Indels 238;	Gaps 2

Oy	ATTGACTAGTAGTTCATTTCCCTGGCAACACTGTAGTAATCCTCGCGTTTTCTGTGGC	651
Db	ATTGACTAGTAGTTCATTTCCCTGGCAACACTGTAGTAATCCTCGCGTTTTCTGTGGC	60
Oy	AATGTTAACCCAGACAAACGGCAACGGGAGGCCCGCGCGCGCGCGCCCAAAGGGCAAGA	657
Db	AATGTTAACCCAGACAAACGGCAACGGGAGGCCCGCGCGCGCGCGCCCAAAGGGCAAGA	120
Oy	CMAACGGGCAACGCCTGCGCGCGCGGAACGGCAAGACAACGGGCAACGGCGCGCGCGGT	663
Db	CMAACGGGCAACGCCTGCGCGCGCGGAACGGGCAAGACAACGGGCAACGGCGCGCGCGGT	180

OY	6637	GGAGTGGAAATTTCGCGCGGGGCAAGGA	CGGATCTTGCGGACGACGCGGGGCGAAGACAG	6639
Db	181	GGAGTGGAAATTTCGCGCGGGGCAAGGA	CGGATCTTGCGGACGACGCGGGGCGAAGACAG	240
OY	6697	CATCCGGGCGATACGGTACAAAGATCAG	CGCGAGCGTGGAGGAGACGGGCGCGGCCGT	6756
Db	241	CATCCGGGCGATACGGTACAAAGATCAG	CGCGAGCGTGGAGGAGACGGGCGCGGCCGT	300
OY	6757	GCTGCGGTGGCCCACGGTGACCGCGT	CGGTATCCCGGCGCTTTCGCGACGGCGGTGAGAGC	6814
Db	301	GCTGCGGTGGCCCACGGTGACCGCGT	CGGTATCCCGGCGCTTTCGCGACGGCGGTGAGAGC	360
OY	6817	CGAAGACGCGGTGCGCGCGCGCTTG	CGCGACCGGCGCATGTTCACTGCTACGCGCGCGCGT	6876
Db	361	CGAAGACGCGGTGCGCGCGCGCTTG	CGCGACCGGCGCATGTTCACTGCTACGCGCGCGCGT	420
OY	6877	CGGCTCTCCCGCGCGCGACGAGGTAC	ATTACAGCTTCAACGGTATGATGCTGAGCAT	6933
Db	421	CGGCTCTCCCGCGCGCGACGAGGTAC	ATTACAGCTTCAACGGTATGATGCTGAGCAT	439
OY	6937	GCATGCGCGGGTTTACTTACGTGCGCG	CGCTGTTCTTCCCGGCGCTTCAAAATTTTA	6999
Db	440	GCATGCGCGGGTTTACTTACGTGCGCG	CGCTGTTCTTCCCGGCGCTTCAAAATTTTA	439
OY	6997	ACCTTCTATAAGTACTTTATAAAACA	AGCGCGCGGTAGACAGCATTTGTCAAGGGC	7056
Db	440	ACCTTCTATAAGTACTTTATAAAACA	AGCGCGCGGTAGACAGCATTTGTCAAGGGC	469
OY	7057	GTGCGCTTACAGCTTATCGGCGCGAG	CGTCTTCTGCAACGCGCGCGCGAAGCTCAAGCGCATC	7116
Db	470	GTGCGCTTACAGCTTATCGGCGCGAG	CGTCTTCTGCAACGCGCGCGCGAAGCTCAAGCGCATC	529
OY	7117	GAAGTGATTAATCCGCGGTGCTGCGC	CGACCTGCGCGCGCAACATATCTGCTTCCCGCGCA	7176
Db	530	GAAGTGATTAATCCGCGGTGCTGCGC	CGACCTGCGCGCGCAACATATCTGCTTCCCGCGCA	589
OY	7177	GCGTATCCAAATTACAGAGCGCGAG	CGGCGATTCAACAAGCTGAGAGTCCGCGCATTTGCAC	7233
Db	590	GCGTATCCAAATTACAGAGCGCGAG	CGGCGATTCAACAAGCTGAGAGTCCGCGCATTTGCAC	649
OY	7237	CTCATCCCGGCAAGAGGGGTGGAGAT	GCATTCGACTGCTGGATCATGCGCGCGACAG	7299
Db	650	CTCATCCCGGCAAGAGGGGTGGAGAT	GCATTCGACTGCTGGATCATGCGCGCGACAG	709
OY	7297	AACACCAACCGGATGCTCATATAACC	CAAAATCCGATCCGCGCGAGGCTTACTCTAC	7356
Db	710	AACACCAACCGGATGCTCATATAACC	CAAAATCCGATCCGCGCGAGGCTTACTCTAC	769
OY	7357	GACCATCTGGCGCAAGTTTGTGATCC	ATGCAATCCTCTGCGCTGTGATCGACCGGTCTGT	7416
Db	770	GACCATCTGGCGCAAGTTTGTGATCC	ATGCAATCCTCTGCGCTGTGATCGACCGGTCTGT	782
OY	7417	TTGAAACATAGATATGAGATTGCGT	TTTGCTATCGTGTGATGATGCTGTTGGTTATC	7476
Db	783	TTGAAACATAGATATGAGATTGCGT	TTTGCTATCGTGTGATGATGCTGTTGGTTATC	782
OY	7477	AGTGTGCGGAGGTGCGCAAGAAAGCT	CCGAAATTTGTGTGATGCTGTGACGAGGTTTACGGCA	7533
Db	783	AGTGTGCGGAGGTGCGCAAGAAAGCT	CCGAAATTTGTGTGATGCTGTGACGAGGTTTACGGCA	842
OY	7537	AACGTGTTCTGGGAGAGCGCCCGT	TTTATCCCGATGGGCGTCTTTGGGCAATTCGCCCGG	7599
Db	843	AACGTGTTCTGGGAGAGCGCCCGT	TTTATCCCGATGGGCGTCTTTGGGCAATTCGCCCGG	902
OY	7597	TCTTTGTCATTGATCTCTGTCCAAG	TGCTGAGATGCTGATGCGCATTTGAGATGGG	7656
Db	903	TCTTTGTCATTGATCTCTGTCCAAG	TGCTGAGATGCTGATGCGCATTTGAGATGGG	962
OY	7657	TGGGGGTGTACAGACCCCAAAAGAT	TTTATGAGAAAATTAAGAGTACTCT	7703
Db	963	TGGGGGTGTACAGACCCCAAAAGAT	TTTATGAGAAAATTAAGAGTACTCT	1009

RESULT 6
AP004116/c 126315 bp DNA Linear HTG 21-MAR-2002
LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone
DEFINITION OJ1121_D06, *** SEQUENCING IN PROGRESS ***.
ACCESSION AP004116
VERSION AP004116.1 GI:15375102
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretioideae; Oryzae; Oryza.
REFERENCE 1
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OJ1121_D06
JOURNAL Published Only in Database (2001)
REFERENCE 2
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Submitted (29-AUG-2001) Takuji Sasaki, National Institute of
Direct Submission
JOURNAL Agrobiological Sciences, Rice Genome Research Program; Kamondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rtp.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and RGP-Japan sequencing data.
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the pieces between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. 126315
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OJ1121_D06"
ORIGIN
Query Match 3.3%; Score 366.8; DB 2; Length 126315;
Best Local Similarity 55.7%; Pred. No. 2.2e-37;
Matches 1082; Conservative 0; Mismatches 692; Indels 170; Gaps 13;
Oy 7008 GTACTTATAAACAAGCGCGCTAGACGACACTGTGACAGGCGCCCTACAA 7067
Db |||||
Oy 24310 GTGCTCTTCTTTGAAATATATGCTGTGGAGATCATTTGTCAGCACTCCCAACAA 24251
Db |||||
Oy 7068 GCTATCGGCGCAGCAGCTCTCTCACCGCGCGGAGAACTCAGCGAGTCGATCAAT 7127
Db |||||
Oy 24250 GCTATCTTCTGATGACATCTTCTTAACCGCTGAGGAACTCAGCGATCGATAT 24191
Db |||||
Oy 7128 CCGCGTGTGCGCCAGACTGCGCGCGCCACATCTGCTTCCCGCGAGGCTATCCAA 7187
Db |||||
Oy 24190 CTCAATCTTGGCCCAACT--GGCACAACAATATGCTTCTTGAACAGGCTACCCAA 24134
Db |||||
Oy 7188 TTACAGGCGGAGCGGATTCACAGCTGAGGCTCGGACCTTCAGCCATCCCGCA 7247
Db |||||
Oy 24133 CTATGAAGCTGACCGCGGTCAACAACCTTGAAGTGTGATCTTATCTTCTGA 24074
Db |||||
Oy 7248 CAAGGGGTGGAGATTCGACATGCTGCTGGAATTCATCGCCGCAAGAACACACCGC 7307
Db |||||
Oy 24073 GAGGGGTGGAGATTCGACATGCTGCTGGAATTCATCGCCGCAAGAACACATCTGC 24014
Db |||||
Oy 7308 GATGCTATCATTAACCAACATCCGTGCGGAGGCTTACTCTACGACATCTGGC 7367

Db 24013 GATAGTCATATAAACCAATATATCATCGCGGAATGTGTACCTTAGAGCATTTATC 23954
Oy 7368 CAAGGTTTGATCATCATGATCTCTGCGCTGTGATTCGACCGGCTGTTGAACATAGT 7427
Db 23953 CAAGGTTTGATCATCATGATCTCTGCGCTGTGATTCGACCGGCTGTTGAACATAGT 23894
Oy 7428 ATATGATTCGCTGTTGCTAATCGTGTGATGATGATGCTGTTGATTCAGTCCGAG 7487
Db 23893 TTA-----ATGTGATATGATTTATTTATTTATTTAGTGGGAGAG 23855
Oy 7488 GTGGCAAGAGCTCGGAATTTGTTGATGCTGACGAGGTTTACGCAACTGCTTCTG 7547
Db 23854 GTAGCAAGAGAGCTCGGAATTTGTTGATGCTGACGAGGTTTATGAGGTTGTTT 23795
Oy 7548 GAGAGGCGCGGCTTATTCGATGGGCGCTTTGGGCAATGCGCGCTGTGCAT 7607
Db 23794 GAGAGTTCCTCATTTGTCCTCAATGGGTTGCTTTGGGCAATGCTATCAATATTAACCA 23735
Oy 7608 GATCTCTGTCGAAGCTGATGATGATGCTGATGCGACTTGATGATGGTGGCGGTGAC 7667
Db 23734 GATCTCTGTCGAAGCTGATGATGATGCTGATGCGACTTGATGATGGTGGCGGTGAC 23675
Oy 7668 GATCTCTGTCGAAGCTGATGATGATGCTGATGCGACTTGATGATGGTGGCGGTGAC 7717
Db 23674 GATCTCTGTCGAAGCTGATGATGATGCTGATGCGACTTGATGATGGTGGCGGTGAC 23615
Oy 7718 TTTCTGATCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7777
Db 23614 TTTCTGATCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 23558
Oy 7778 ATTCAGATCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7837
Db 23557 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23498
Oy 7838 GTTCAGGTTAGTCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7878
Db 23497 ATTCAGGTTAGTCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 23438
Oy 7879 CCTGTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7938
Db 23437 ATCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23378
Oy 7939 GAATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7998
Db 23377 GATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23318
Oy 7999 TT-----TGTGATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8052
Db 23317 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 23258
Oy 8053 GAATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8095
Db 23257 AAACATAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23198
Oy 8096 CATATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8155
Db 23197 CTATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23138
Oy 8156 ATGTTGT----- 8162
Db 23137 ATTTGTGTCATCTGAGGTGCGATTAATTAATTAATTAATTAATTAATTAATTAAT 23078
Oy 8163 -----TAAGATTTCTTCACTACACACCTTGTCCGATCTTCAATTTTG 8210
Db 23077 GTTAAGTGAAGTATCAATTTCTTAATCTGAGATTAATTAATTAATTAATTAATTAAT 23018
Oy 8211 ATATCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8270
Db 23017 ATATATTTGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22959
Oy 8271 ATCAATTTGTTGAGAGCTCTTCTTAATTAATTTTGAAGAACAAAGAGATTTCTTTA 8330

Db	22558	GTGTAATTTCTTAGGAGGCTCTACCGAAATTTCTTAAAGATACCAAGAAAGAAATCTTTA	228939
Qy	8331	AGAGGAAATTTGGTCTCTAAAGGAATCATCGAGATAGTTTATAGGAAATTAAGGAAA	8390
Db	22898	AGAGAGATTAATGAAATTTCTTACGGAAACATCAAGATTTTGTCTTAAGAGAAATTAAGGATA	228393
Qy	8391	ACAAATTAATTAAGTGCTCCAAAGCCAAAGAGATCGATGTTGTAATGTAAGTAAAG	8450
Db	22838	TTAAATCATCATCTTGCTCTCAAGCCGAGAGATCATGTTTGATGAGTAAAGATTTTC	227797
Qy	8451	CATAGACTTACTTTTAAAGTTAATCTGGAGATCTCAGTGATCCAAACAAATCAATC	8510
Db	22778	CTCTTAATCATTTTATCATTAAGCTGTACT-----AGATCAATTTATCAACATC	22722
Qy	8511	AAAAATTAATTAATGTTTGGCTAATGAGATCTTTTGAAGATGCATGATTTGAAGAAATATG	8570
Db	22726	AAAAAATTAATTAATGCTAAT--ATTTAATTAATGTTGGACCTTAACATTAATTCAGATGC	22668
Qy	8571	AAGAGATGGAATTAATTTTAGAGCTAATCTCTGATATCATTTTGTCCATTTTTTGT	8630
Db	22668	ATGAGA-----AATTTTGGGTATTTGTGTGAAGATTAATGCTAATTAATTAAGAGAT	22617
Qy	8631	ATTACTGTAATGATGTAACACTCAAAATCATATTACAAAAGTTTCCCATTTTATGTA	8690
Db	22616	TGTAGTTTGCATGTTAAACATCAATTTTCTTT-----GTGCTACAAACATACTTGTC	22564
Qy	8691	AGATTAAGCTTCTCTTCTATTAACCATGATTAATTAAGCTCATGTAACAGTCAAACTTAATCT	8750
Db	22563	ACTTAAGCTTCACTTCCAGTAATCTTACATTGAATTCATTAACAGGTAAATTTGAATC	22504
Qy	8751	TACATCTTTTGGAGAGATCATGACGACATAAATTTTGTCTGCAAGCTGCCAAGGAAG	8810
Db	22503	TATATCTTTTGGAGAGGATCCATGATGATGATGATTTTGTGCCAACTTGCAGAAAGAG	22444
Qy	8811	AATCTGTAATTTTAATGCCAGGTA	8834
Db	22443	AGTGGTGATTTCTTTGCCAGGTA	22420

RESULT 7	
APO05743/c	
LOCUS	AP005743 159075 bp DNA linear PLN 15-UN-2004
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
ACCESSION	PAC clone:P0543Ct1.
VERSION	AP005743
KEYWORDS	AP005743.3 GI:48716873
SOURCE	.
ORGANISM	Oryza sativa (japonica cultivar-group)
TITLE	Oryza sativa (japonica cultivar-group) Eukaryotic Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyte; Magnoliophyta; Liliopsida; Poales; Poaceae; Euhartoidaeae; Oryzaceae; Oryza. 1 Sasaki,T., Matsumoto,T. and Katsuyose,Y. Oryza sativa nippondare(GMs) genomic DNA, chromosome 2, PAC clone:P0543Ct1 Published Only In Database (2002) 2 (bases 1 to 159075) Sasaki,T., Matsumoto,T. and Katsuyose,Y. Direct Submission Submitted (11-Sep-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:takeshi@saffrc.go.jp, url:http://rsg.dna.affrc.go.jp/, Tel:+81-298-38-7441, Fax:+81-298-37-7468) On Jun 14, 2004 this sequence version replaced gi:38564220. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mil.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), Genemark.hmm (http://opal.biology.gatech.edu/genemarf/), GlimmerM (http://www.cigr.org/tadb/glimmerm/glmr.form.html), RceHMM (http://rsg.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/bp.cgi), sim4

```

(http://globin.cse.psu.edu/html/docs/s1m4.html), g9a2
(http://www.tigr.org/software/glimmer/), BLASTn and BLASTx. The
genomic sequence was searched against NCBI NonRedundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DDBJ. Protein homologues of the coding
regions were searched against NCBI NonRedundant Protein database
with BLASTn. ESTs represent the identified cDNA sequences using
BLASTn with the corresponding DDBJ accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTn with the corresponding DDBJ accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from 5' to 3' of the PAC clone.
This sequence of P0543C11 clone has an overlap with P0572A04 clone
(DDBJ: AP005009) at 5' end and an overlap with OSJNB0026D20 (DDBJ:
AP005476) at 3' end. Detailed information on overlap and assembly
quality together with annotation of this entry is available at
http://xrp.dna.affrc.go.jp/GeneSeq.html.
Location/Qualifiers
1. 159075
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0543C11"
4637..4982
/gene="P0543C11.1"
join(<4637..4749,4901..>4982)
/gene="P0543C11.1"
/notes="start and end point are not identified"
join(4637..4749,4901..4982)
/gene="P0543C11.1"
/notes="predicted by Glimmer etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD23570.1"
/db_xref="GI:48716874"
/translation="MTDVTGLTPVMSADSHDSTHLLPAKVEVWVGRLLRRIGTQNK
GGAGIVVVASVYDLPKTOE"
5899..7665
/gene="P0543C11.2"
5899..7665
/gene="P0543C11.2"
/notes="probably inactive due to 5' exon missing in CDS
probably inactive due to including stop codon(s) in CDS
pseudogene, GAG-POL precursor"
pseudo
complement(9157..9750)
/gene="P0543C11.3"
complement(9157..9750)
/gene="P0543C11.3"
/notes="polyprotein-1like"
complement(join(10709..10815,11184..11241))
/gene="P0543C11.4"
complement(join(10709..10815,11184..11241))
/gene="P0543C11.4"
/notes="hypothetical ORF
predicted by GENSCAN
this category is not included in IRSP standard"
11463..12659
/gene="P0543C11.5"
join(<11463..11535,12538..>12659)
/gene="P0543C11.5"

```


CDS
 /note="start and end point are not identified"
 join(11463..11535,12538..12659)
 /gene="P0543C11.5"
 /note="predicted by GENSCAN etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD3571.1"
 /db_xref="GI:48716875"
 /translation="MSMRAGVLLGGSTLPACPGSGLLDGVSEEDIVSLPKLR
 KTGSASEVEAYKPAVKOR"
 12934..122525
 /gene="P0543C11.6"
 12934..122525
 /gene="P0543C11.6"
 /note="probably inactive due to including stop codon(s) in
 CDS
 pseudo gene, polypeptide"
 /pseudo
 complement(23001..24433)
 /gene="P0543C11.7"
 complement(join(23001..23488,23511..24433))
 /gene="P0543C11.7"
 /note="supported by full-length cDNA(s) : AK119914"
 complement(23520..23969)
 /gene="P0543C11.7"
 /note="retrotransposon protein-like"
 complement(25886..26575)
 /gene="P0543C11.8"
 complement(25886..26575)
 /gene="P0543C11.8"
 /note="probably inactive due to no termination codon in
 CDS
 pseudogene, orf3 of RIR2"
 /pseudo
 complement(join(27629..27723,28537..28651))
 /gene="P0543C11.9"
 complement(join(27629..27723,28537..28651))
 /gene="P0543C11.9"
 /note="hypothetical ORF
 predicted by GENSCAN
 this category is not included in IRGSP standard"
 28948..28562
 /gene="P0543C11.10"
 join(<28948..28960,29390..29562)
 /gene="P0543C11.10"
 /note="start and end point are not identified"
 join(28948..28960,29390..29562)
 /gene="P0543C11.10"
 /note="predicted by GENSCAN etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD3572.1"
 /db_xref="GI:48716876"
 /translation="MTPTAAGRGSPAGAPAAAGCGPVAAGLAARAGRPAWR
 LRHVAFFFGELVIRI"
 complement(31692..38144)
 /gene="P0543C11.11"
 complement(join(31692..31946,32033..32113,32444..32602,
 32965..33084,33640..33720,33801..33869,34425..34538,
 35083..35136,36066..36199,37805..38144))
 /gene="P0543C11.11"
 /note="supported by full-length cDNA(s) : AK102077"
 complement(join(31872..31946,32033..32113,32444..32602,
 32965..33084,33640..33720,33801..33869,34425..34538,
 35083..35136,36066..36199,37805..38060))
 /gene="P0543C11.11"
 /note="contains EST(s) : C98634(E0488),A0162173(E0488)
 contains full-length cDNA(s) : AK102077"
 /codon_start=1
 /product="unknown protein"
 /protein_id="BAD3573.1"
 /db_xref="GI:48716877"
 /translation="MAAAEAVALFLHHHLAFGLRIISVSVPSPRRSAGVALLA

VVVVALLATTAATGATIESRPADLRSEVDRLRLVLESLLEENTKLKSKANN
 LEENSLNIGMEHDIILNNKYEETKSSQSPENRYALREVDLLEVMKRN
 ADSIESLNGAEKRVESLSFVKMEKQVLAEMIOLEAVALTRMVKHRSR
 LBETATKMRGKDLVTKFRLNIGTFLMGYSYTSKCSHYTKRSTQAMNRPYHVS
 RFCKALCGHIRVDKPNVFLGGSISRCSISAPYQKIFMLAONFHHKVOIFLQD
 AMNSNSYSRGPAVEITPCLAFVVISPMWIMFLVSTRGSKK"
 complement(39218..39454)
 /gene="P0543C11.12"
 complement(39218..39454)
 /gene="P0543C11.12"
 /note="hypothetical ORF
 predicted by Glimmer"
 this category is not included in IRGSP standard"
 40782..46296
 /gene="P0543C11.13"
 join(40782..41047,44517..44580,45626..45698,45821..45876,
 45968..46296)
 /gene="P0543C11.13"
 /note="supported by full-length cDNA(s) : AK066730"
 join(44525..44580,45626..45698,45821..45876,45968..46094)
 /gene="P0543C11.13"
 /note="contains EST(s) : AU031536(E61810),AU031537(E61810)
 Query Match 3.3%; Score 366.8; DB 8; Length 159075;
 Best Local Similarity 55.7%; Pred. No. 2.3e-37;
 Matches 1082; Conservative 0; Mismatches 692; Indels 170; Gaps 13;
 CDS
 7008 GTACCTATTAATAAACAAGCCCGCTGACAGAGCACTGTCTACAGGCGTCCATCAA
 96676 GTCCCTTTCTTTGAATATAGCTGTGTGAGATCTTGTCTACGCACTCCCATTA
 7068 GCTATCGCCGACGAGCTCTCTCTACCGCCGCGAAGTCAAGCCATCAATAT
 96616 GCTATCTTGATGATGATCTTCTTAACGCTGAGAACTCAGCCATGAGTCTGAT
 7128 CCCGTCGTCGACGAGCTGCGGCGCAATATCTCTTCCCGGCGAGCTATCCAA
 96556 CTCATCTCTGCCCACCT---GGCAAAACATATGCTTCTTACGACGAGTACCCAA
 7188 TTAAGAGGCGGCGGCAATCAAGCTGAGGCTCCGCACTTGACCTCAATCCCGA
 96499 CTATGAGCTCGAGCGCGCTTCAACACCTGAAGTTCGCTTGTATTTCTGA
 7248 CAAGGCTGAGAGTATGATCACTGCTGATCATCATGCGGAGAACACACCGC
 96439 GAAGGCTGAGAGTATGATCACTTAACTCCATGATATATTGGGCAAGAACTATGC
 7308 GATGTCATCATTAACCAACATATCGTGGGAGAGCTTTACTCTTACATCTGCGC
 96379 GATAGTCATCATTAACCAATATCATGCGGAGATGTGTATCATTAACGATTTATC
 7368 CAAGTTTTCATCAATGATCTCTGCTGATGATGACCGGCTGTTGAACATAGT
 96319 CAAGTTTTCATGCTGTTATTTAATTTAATTAATTAATTTATGCGCACTTA
 7428 ATATGATTCGCTTCTATATCTGTGATGATGATGCTGTTGTTATCAAGTGGAG
 96259 TTA-----ATGATCATATGATTTATTGTTATTTAGTGGGAGAG 96221
 7488 GTGCGAAGAGTCTGGAATATTTGATGATGCTGATGAGAGTTTACGCAACTGTTCTG
 96220 GTAGCAAGAGAGTCTGGAATATTTGATGATGATGATGATGATGATGATTTGTTT
 7548 GGCAGCGCCCGTTATTCGATGAGGCGCTTTGGGACATTTGCCGCTTGTCCATT
 96160 GGGAGTTCCCATTTTTCATATGAGTGTGCTTTGGGACATCTATCAATATTAACATA
 7608 GATCTCTGTCAAGTGTGATGATGATGCTGATGAGCGCACTTGGATGGTGGCGGTAC
 96100 GATTCCTATCAAGAGTGTGATGATGCGGAGTGGCACTTGGTGGTGAAGATATGT
 7668 GACCCCAAAAGTTTATAGAAAATAAGTAGTCTT-----TAGTCCCATCA 7717

Db	96040	GACCCCAAGAAAGACTACACAAAGAAACCAAGTAATAATATATTAACTGTTTGANA	95981
Qy	7718	TTCTTTCATATGCTACTGTGGGATTAAGTATTTTGGCTAAATTTGATCGCTTTGTTT	7777
Db	95980	CTTTTGTGTAATGTAATATTTGAAATATATATTTTCTTAAGTGATACAC--TTCT	959244
Qy	7778	ATTGAGATCTCTACGTCATATTAAGAAATTTACCTTAATGTCTCAACGAGACCAGAACCTTC	7837
Db	95923	TTTGAGATTGCAACTTAATATTACATAATTTCTTAATGTTCAACTGATCCAGAACCTTC	95866
Qy	7838	GTTCAAGTTAGCTTTGGTCTTGCC-----TAATTTGCTAGT	7878
Db	95863	ATTGAGGTTAGTGTTTTTCCTTAACAGCCCAATGTTGGGTTATATATTTGGTCATTT	95804
Qy	7879	CCCGTGTCGATGCTCAAAATGACCGGCTCAAGTATGATATAGAGTTTGTTAAGTGT	7938
Db	95803	ATCTGCTTATATGATGTTTGGGACTATGTTGTGGTCTTCGGAAATGGATG	957444
Qy	7939	GAATGTGGAATGCCAATGATGGAAGAAAGATACATATTTTATGATCAATCCCTTTG	7998
Db	95743	GCTGTCAAAATGACAGTGGGAGTGTACACTGTTGTGATGCTGAGAAAACACTAGTGT	9568
Qy	7999	TT-----TGTTGATTCATTAATAATTAATTAACCAAGCCGAAACCACTATTGAAT	8052
Db	95683	TTAAATATATCTATCGAATCTTACCAAAATTAATAGTACTTATGAATTCAGAGTTATATATC	956244
Qy	8053	GAACATATTTTCTTGAAGAAATATACA-----TTGATTTTGGACATGC	8095
Db	95623	AAACATATAGTTCAATCAAAAGTACAAAAGAAAGTTAATTCAGTTTATTTATGTTCACTTCAAT	95566
Qy	8096	CATATCTTTTCGATCAAGATGCAATATATTAATACTTGATTTGATCAAGATATACC	8155
Db	95563	CCATATTCCTGTAAACATATATATATGTTTAAAAATTAATTAAGAAATGAGGTTGT	955044
Qy	8156	ATGTTGT-----	8162
Db	95503	ATTTTGTCTCACTCTGAGTGGGTATTTTAAAGTATATGTAAGTTCAATCTTAC	954444
Qy	8163	-----TAAAGATTCCTTAACCTTACAACACCTGTCGAGTCTTCATATTTTG	8210
Db	95443	GTTAAAGTAAAGCATACATCTTAATTAATGATTAATTATTTGATCTTGCTTTAAAT	953844
Qy	8211	ATATCCTTGACATATGTCTCTTAATGATTCACACACTTAATTAATGATTTTGTCT	8270
Db	95383	ATTATATTTGCGTCGTGATTCATTTTATCTCAAAATTAATGAATTTGCAATGTT--TGTT	953255
Qy	8271	ATCAAAATTTGTAGGAAGCTCTTCTTAATTTCTTGGAACAACAAGACAGATTTCTTA	8330
Db	95324	GTTGAATTTCTTAGGAGCTCTACCGAATTTCTTAAGATACCAAGAAAGATTTCTTA	952655
Qy	8331	AGAGGATTTTGTCTACTAAAGGAATCATCAGAGATATGTTATTAAGGAATTAAGAAA	8390
Db	95264	AGAGGATTAATTTGATTTGCTTACGGAACAATCAATATTTGCTTAAGAAATTAAGATA	952050
Qy	8391	ACAAATATTTAGTGTCCTCAAGCCGAAAGATCGATGTTGTAATGTAAGCTAAG	8450
Db	95204	TTAAATCACTACTTGTCCTCAAGCCGAAAGATCAATGTTTGTGATGTAAGATTTTC	951450
Qy	8451	CATAGACTTACTTTTAAGGTTATCTGGATCTCAGTGATCCAAACAACATCAATC	8510
Db	95144	CTCTTAATATCATTTTATCATTAAGGCTACT-----AGCATCAATTTATCACATC	950933
Qy	8511	AAAAATATATTAATGTTTGTCTATGAGATCTTTTGAAGTCAATGATTTGAAGATATAG	8570
Db	95092	AAAAATATATTAATGCTAT--ATTTAATTAATGTTGGCACTTAACAACCTTAATTTCAAGATC	950355
Qy	8571	AAGAGATGGAATTAATTTTAGGACTAATCTTCGTAATCATTTTGTCAATTTTGTGT	8630
Db	95034	ATGAGA-----AATTTTGGGATATGTGTTGAAGATTAATGCTATTAATTAAGAAAT	949833
Qy	8631	ATTATCTGTAATTTGTATACACTCAAAATCATATTAACAAGAAAGTTTCTCCCATTTTATGTA	8690
Db	94982	TGTAGTTTGCACTGTAATCAATATTTTCTTTT-----GTGCTACAAATACCTTTGTC	949300

QY	8691	AGATTGACCTTCCTTTATATACAGTATGATTTAACTTCATGTAACAGGTCAAACTAACT	8750
DB	94299	ACTTAAGCTTACTCTTCAGTATATCTTACATTTGAATCTATTAACAGGTGAATTTGAAC	9487
QY	8751	TACATCTTTTGGAGAGATTCATGACGACATMAATTTTGTCTGCAAGCTCGCAAGAG	8810
DB	94869	TATATCTTTTGGAGAGATTCATGATGATGATTTTGTGCGCACTTCGAAAGAG	9481
QY	8811	AATCTGTAATTTATATGCCAGTA	8834
DB	94809	AGTGGTGATTTCTTGGCCAGSTA	94786

LOCUS	BT009504	1035 bp	mRNA	linear	PLN 20-JUN-2003
DEFINITION	Triticum aestivum clone wrl.pK0085.h9:fls, full insert mRNA sequence.				
ACCESSION	BT009504				
KEYWORDS	BT009504.1	GI:32129055			
SOURCE	FLI CDNA.				
ORGANISM	Triticum aestivum (bread wheat)				
REFERENCE	Triticum aestivum				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.				
TITLE	1 (base 1 to 1035)				
JOURNAL	Tingey,S.V., Wolters,P., Powell,M., Dolan,M., Miao,G.-H., Caraher,N.R., Hanafey,M.K. and Hainey,C.F.				
FEATURES	Direct Submission				
SOURCE	Submitted (20-JUN-2003) Crop Genetics, E. I. Dupont de Nemours and Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA				

ORIGIN	Location/Qualifiers
1..1035	/organism="Triticum aestivum"
/mol_type="mRNA"	
/db_xref="taxon:4565"	
/clone="wrl.pK0085.h9:fls"	

Query Match	3.0%; Score 326; DB 8; Length 1035;	
Best Local Similarity	75.5%; Pred. No. 3,1e-33;	
Matches	465; Conservative 0; Mismatches 55; Indels 96; Gaps 1;	
QY	1554 GCGGAGCCCGAGGCGATGAGGTCAATATCCGGTGTGCGCCAGACCGCGGCCACACA	16133
DB	1 GCACGACCCAAAGCATTTGAGGTCAATATCCAGTTCTGCGCCAAACCGCGGCCACACA	60
QY	1614 TTCTGCTCCCGAGGCGGCTACCCAAATACGAGGGCGCGCGCTTCAACAGGGTGG	16733
DB	61 TTCTGCTCCCGAGGCGGCTATCAATCAATACGAGGGCGCGCGCTTCAACAGGGTGG	120
QY	1674 AGGTCCGGAATTTGACCTCATCCCGCAAGAGGGGTGGAGATGACATGACCTGCTGG	17333
DB	121 AGGTTCGACATTTGACCTCATCCCGAGAGAGGGGTGGAGATGACATGACCTGCTGG	180
QY	1734 AATTCATGCGCGACAGAACACACACCGCATGTGATCATTAACCCCAACACCCGTGGC	17933
DB	181 AATTCATGCGCGAGAACAAACACACCTCGATGTGATCATTAACCCCAACATCCATGTG	240
QY	1794 GCAGCGTTTACTCTTAGACCATATGTCGAAGTTTCAATCTTTGCTGTGATATAT	18533
DB	241 GCAGCGTTTACTCTTAGACCATATTTAGCCAAAGTT-----	275
QY	1854 GGATTCAGTTCAATGACCTGCTGAATCTTTTGGCAATGCAATGACTGACATGTGC	19133
DB	276 -----	275
QY	1914 TCAATTAAGTTCGCGAGAGTGGCGAAAGGCTTCGGAATATTTGGATTTGTCAGAGGTAT	19733
DB	276 -----GCAGAGGTGGCAAGAGCTTCGGAATATTTGGATTTGTCAGAGAGGTAT	324

QY	1974	ACGGACACTGGTCTCGGAGACGCGCCCGTTCATCCCAATGGAGATGTTGGACATCA	2033
Db	325	ACGGACACTGGTCTTTGGCTGCGCCCGCTTCATCCCAATGGAGATGTTGGACACTCA	384
QY	2034	CCCTGTCTGTCCATAGAGGTCTGTCCAAAGTCATGATAGTGCCTGGATGCGGCTTG	20939
Db	385	CCCTGTGTGTGCCATAGAGGTCTGTCCAAAGTCATGATAGTGCCTGGATGCGGACTTG	444
QY	2094	GATGGGTAGCGGTGTACACACCCCAAGATCTTACAGAAACTAGACTTAATCTC	21533
Db	445	GATGGGTAGCGGTGTACACACCCCAAGATCTTATAGAGAAACTAGACTTCACTCA	504
QY	2154	TATACATCTTTTCA	2169
Db	505	TTAGCAATTAACCTTA	520
RESULT 9			
AP005310		142166 bp	DNA
LOCUS			linear
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,		PLN 15-JUN-2004
ACCESSION	AP005310		
VERSION	AP005310.3		GI:48716578
KEYWORDS			
SOURCE			
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Sasaki, T., Matsumoto, T. and Katayose, Y.		
AUTHORS	Oryza sativa nipponbare (GM3) genomic DNA, chromosome 2, PAC		
TITLE	clone, pO477B05		
JOURNAL	Published only in Database (2002)		
REFERENCE	2 (bases 1 to 142166)		
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-MAY-2002) Takui Sasaki, National Institute of		
COMMENT	Agrobiological Sciences, Rice Genome Research Program, Kamondai		
	2-1-2, Tsukuba, Ibaraki 305-8602, Japan		
	(E-mail: tsasakik@nias.affrc.go.jp, URL: http://xgp.dna.affrc.go.jp/,		
	Tel: 81-298-38-7441, Fax: 81-298-38-7468)		
	On Jun 14, 2004 this sequence version replaced gi:37497098.		
	Genes were predicted from the integrated results of the following:		
	GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), GENESH		
	(http://www.softberry.com/), GeneMark.hmm		
	(http://opal.biology.gatech.edu/GeneMark/), GlimmerM		
	(http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM		
	(http://xgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor		
	(http://bioinformatics.lastate.edu/cgi-bin/sp.cgi), sim4		
	(http://globin.cse.psu.edu/heml/docs/sim4.html), gap2		
	(http://www.tigr.org/software/glimmerm/), BLASTN and BLASTX. The		
	genomic sequence was searched against NCBI Nonredundant Protein		
	database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA		
	sequence database at RGP or DBP. Protein homologies of the coding		
	regions were searched against NCBI Nonredundant Protein database		
	with BLASTP. ESTs represent the identified cDNA sequences using		
	BLASTN with the corresponding DBJ accession no. and RGP clone ID.		
	Full-length cDNAs represent the identified cDNA sequences using		
	BLASTN with the corresponding DBJ accession no.		
	A gene with identity or significant homology to a protein is		
	classified based on the protein name to indicate the homology level		
	such as same name, 'putative-' and '-like protein'. A gene without		
	significant homology to any protein but with full-length cDNA or		
	EST homology (covering almost the entire length of partial		
	sequence) is classified as an 'unknown' protein. A gene predicted		
	by two or more gene prediction programs is classified as a		
	'hypothetical' protein according to IRSP standard. A gene		
	predicted by a single gene prediction program is also classified as		
	a 'probable 'hypothetical' protein and is included as a		
	miscellaneous feature of the sequence.		

The orientation of the sequence is from Spe I to T7 of the PAC clone. This sequence of P0477B05 clone has an overlap with OSUNBa0010K08 (DDBJ: AP005532) clone at 5' end and with P05172A04 (DDBJ: AP005009) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://kcp.dna.affrc.go.jp/GenomeSeq.html>.

gene
/organism="Oryza sativa [Japonica cultivar-group]"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0477B05"
439. .5934
/gene="P0477B05.1"
join(439. .549,1811. .1882,2141. .2230,2346. .2388,
2631. .2737,4724. .4819,5249. .5439,5539. .5608,5827. .5934)
/gene="P0477B05.1"
/note="start and end point are not identified"
join(439. .549,1811. .1882,2141. .2230,2346. .2388,2631. .2737
4724. .4819,5249. .5439,5539. .5608,5827. .5934)
/gene="P0477B05.1"
/product="putative cell differentiation protein Rcdip"
/protein_id="BAD23249.1"
/db_xref="GI:48716579"
/translation="NSTGIQVNPPTDEIAQRVLVDLPDRLGHALAELAKEMFPD
LAPLMSGCTIVLQELIVVPVSPPTLSLASSRGNALALQCAVSHETRE
FLKAIPLVLCSELENTSKTRPEYRLTSLGIVLAKVEGLEVINFLQYEPFLC
FLHMAAGSELSTKVAATFLEKTLVDAGAGYCATDRFPAVTAAGWTSMDKES
PLRLKIKHICYLITDNPRLGLEALQCLPTLTIDGTFFNNLTKDPTQWLQELTAVY
RSGKMGFLPPPEGMHRY"
complement(join(8981. .9460,9615. .9812,9980. .11942,
13202. .14057,14312. .14637))
/gene="P0477B05.2-1"
complement(join(8981. .9460,9615. .9812,9980. .11942,
13202. .14057,14312. .14637))
/gene="P0477B05.2-1"
/note="supported by full-length cDNA(s) : AK067818"
complement(join(9218. .9460,9615. .9812,9980. .11942,
13202. .14016))
/gene="P0477B05.2-1"
/note="contains full-length cDNA(s) : AK067818"
/product="putative LZ-NBS-LRR class RGA"
/protein_id="BAD23250.1"
/db_xref="GI:48716580"
/translation="MAVVAAGMTALIAKLTLLMDQYKLQNSARGDIIPKAELESIO
AALEKLEFVQVTDQYKIEWERDVQESYMEDIIDFMVHVEHTLAKPHELGFMFG
SLSLRAVAVATDIDKIRRLVNETSARBRVYGVITAAKETAIDPLIGLY
GSATKLGISGPEELTKLMDSKGSKKVLIVSGVQIGKTLTANVYVOOLGO
PECHAVSVSLKRPDLKVVSLILRQSEGGYAWTEWCAQEIITKIDETKEKRYLIV
IDIDWEKSWECIECALIENDRGSRLITTSRYVDAATPCSSSEVDHTVYKLOPISDWS
KGLFKRIIFYCDGCEPLEKIDISEKTLRCGEVPLAIIITGSLARPOLXOMDRH
NIIGGELGSEVVENNRHLISISYDLPAHLACFELYITPEDVNIQKODLIRWIS
BEFIIQEDMDTLHEGKYNELINRSMIDPAYIDSGRACVACVHMDVLTLSNE
TNPVLSGGQCPVTHNPRARLCLONISYDHTIRQEKMGAPRSLIVPHTNLLPS
LSRPLIRLVDLEGODLKNQIEGSDLPFHRSLVLTXTNIGSLKPKIGNLSCLHTL
DIRHTITLPSVTHLRVLRLVDLDAVSLPDGIGKPCSLQISLVGSKSNPLKE
IGSLTELRLLISSTGAWHDSYENTLIDSLCNHICDLYHGKTLSTEPINFS
PYYLVISGQGLIIPWINSLSISTDGLLVNRODLSLGAQSLCURLVWF
KLEPRLLVGTGHEKHPMLAEPSTTNAGGLIFSOYSMRLELTLAPNVRERKYPDI
GLEHLSLKDVARIIDCRDSIFEVQNDAAIIRAPLVNNSNPKVHVIHYEMLDE
VKIQGTSEBEKSLIENRIGSPWNGNGVYTCDIAPRRLSEYITCSGIIIDALAFSLD
KQGHETHTWGGGLGSGVQLIDDEREFLEMEYIKVGPNNVLSAETLSLTLVNVSYS
GFGPQPGPPTPRKNSCIYFGRSGTYDAVGVYHPM"
complement(10700. .15356)
/gene="P0477B05.2-2"
complement(10700. .15356)
/note="contains full-length cDNA(s) : AK101276"
non-coding transcript

/gene="P0477B05.8"

Query Match	2.5k	Score 273.8	DB 8	Length 142166
Best Local Similarity	65.9%	Prod. No. 2e-25		
Matches 446	Conservative	0	Mismatches 222	Indels 9
			Gaps 3	
Qy	1480	CAGCGCGGTGACAGAGCACTGTCGCAAGGGCGTCCTTACATGCTATCGCGCAGACGT	1539	
Db	40983	CAGGCGGTGCGCGCGCACCTGTGTGCGGAGACTCCCTAAGCGCTTTCGCGCGGAGCT	4104	
Qy	1540	CTTCTCAACCGCGCGGACCCAGGGATTCAGAGTCAATATCCCGGTGCTGCGCCAGAC	1599	
Db	41043	GGTCTCAACCGCGCGGTCAACCAACGCGCTGAGATCATATGTCCTGCTCG--GTC	4109	
Qy	1600	CGCGCGGCAACATTTGCTCCCGCAGGCGAGGTACCCAACTACAGAGGCGCGCGCG	1659	
Db	41100	GCGGCGGCAACGTCTGCTCCCGCGGCCCGCTTACCCGTGTACCGCTGCGCGCGCG	4115	
Qy	1660	GTTCAACAGCGGTGAGGTCCGGCATTTTCACTTCATCCCGACAGAGGGGTGGAGATCGA	1719	
Db	41160	CTTGAAGCGGCTCGAGTTCGCGCACTTCGACCTTCCTCCCGACAGGAGTGGAGTCTGA	4121	
Qy	1720	CATGACCTCGCTGAATTCATCGCCGACAGAACACACCGCCATGTGATCAATAAACCC	1779	
Db	41220	CTTCGCGGCGTCAAGGCTCCCTGCGCAAGCCAAACACCGTCGCGCATGTGATGCAACCC	4127	
Qy	1780	CAACAAACCGTGCAGCGCGCTTTACTCTTACAGACATCTGTCTCAAGGTTTCAATCTTT	1839	
Db	41280	CAACAAACCTCGCGGCTCGCTTACTCTCCGAGACACCTGCGCAAGGTTTGAATTA	4133	
Qy	1840	GCTTGTCTGA--ATATGATTTACGTTCAATGATGACCTGCTGAATCTTTTGGCAATGCCA	1897	
Db	41340	ATACATCTGTATCAAAATCAATTCGATGAGGATGATATACAATTCGTAATCATTAACCGGA	4139	
Qy	1898	TACTGACTGATGTTGCTCAATTAAGTTCGCGGAGGTGCGAAAGGCTCGGAATATTGCTG	1957	
Db	41400	TTAAT-----ATCTGTACTACTAGATGTCAGAGACGCGAGAGCTGGGGTAAATGCTG	4155	
Qy	1958	ATTGCTGACGAGGTATACGGCAAGCTGTTCTGCGGACGCGCCCGTTTCATCCCAATGGGA	2017	
Db	41456	ATCAGCAGCAGAGGTATGACACCACTTGCGCTTGGGAGCAAGCGCTTCGTGCCATGGGG	4151	
Qy	2018	GTTGTTTGGGACATCAACCCCTGTGCTGTCCATGAGGCTCTGTCCTCAAGTACATGATGCTG	2077	
Db	41516	GTTTTCGGCGACGTCGCGCGGTATATACCTGTGGCGGACATCTCAAGCGGTGATGCTG	4157	
Qy	2078	CTTGATATGCGCGCTTGATGAGTGGTATGACGCGGTATACGCCAGAAAGATCTTACAGAAACT	2137	
Db	41576	CCGGGCTGGCGCTCGGCTGGATGGCGGCAACGATCCCAAGGAATCTTCAGAAACAG	4153	
Qy	2138	AAGGTACTTAAATCTCT 2154		
Db	41636	AAGGTCTTAATTGCT 41652		

RESULT 10
AP005532
LOCUS 184760 bp DNA linear PLN 15-JUN-2004
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
BAC clone:OSUNBA0010K08.
ACCESSION AP005532
VERSION AP005532.3 GI:48716663
KEYWORDS Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Euhartophyceae; Oryzae; Oryza.

1
Sasaki, T., Matsumoto, T. and Katayose, Y.
Oryza sativa nipponbare (Gh3) genomic DNA, chromosome 2, BAC
clone:OSUNBA0010K08
Published Only in Database (2002)

REFERENCE 2 (bases 1 to 184760)
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUL-2002) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL: http://rsgp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 On Jun 14, 2004 this sequence version replaced gi:34850228.
 Genes were predicted from the integrated results of the following:
 GENSCAN (http://CCR-081.mlt.edu/GENSCAN.html), FGENESH
 (http://www.softberry.com/), GeneMark.hmm
 (http://opal.biology.gatech.edu/GeneMark/), Glimmer
 (http://www.tigr.org/cdb/glimmer/glimr_form.html), RiceHMM
 (http://rsgp.dna.affrc.go.jp/RiceHMM/) SplinePredictor
 (http://bioinformatics.lastate.edu/cgi-bin/sp.cgi), sim4
 (http://globin.cse.psu.edu/html/doc/sim4.html), gap2
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI NonRedundant Protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 sequence database at RGP or DDBJ. Protein homologues of the coding
 regions were searched against NCBI NonRedundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTX with the corresponding DDBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as
 a miscellaneous feature of the sequence.
 The orientation of the sequence is from -21M13 to M13rev of the BAC
 clone. This sequence of OSJNBa0010K08 clone has an overlap with
 OSJNBa0030M21 (DDBJ: AP005534) clone at 5' end and with P0477805
 (DDBJ: AP005310) clone at 3' end. Detailed information on overlap
 and assembly quality together with annotation of this entry is
 available at http://rsgp.dna.affrc.go.jp/Genomeseg.html.
 Location/Qualifiers
 1. 184760
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultur="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OSJNBa0010K08"
 /complement(2176..2664)
 /gene="OSJNBa0010K08.1"
 /complement(2176..2664)
 /note="OSJNBa0010K08.1"
 /note="probably inactive due to 3' exon missing in CDS
 probably inactive due to 5' exon missing in CDS
 pseudogene, gag-pol polypeptide"
 /pseudo
 /complement(join(2833..2867,3889..4091,5333..5357))
 /gene="OSJNBa0010K08.2"
 /complement(join(2833..2867,3889..4091,5333..5357))
 /gene="OSJNBa0010K08.2"
 /note="OSJNBa0010K08.2"
 /note="polyprotein-like"
 /complement(join(9450..9793,9884..9901,9940..9949))
 /gene="OSJNBa0010K08.3"
 /complement(join(9450..9793,9884..9901,9940..9949))
 /gene="OSJNBa0010K08.3"
 /note="hypothetical ORF
 predicted by Glimmer
 this category is not included in IRGSP standard"
 13662..14499
 /gene="OSJNBa0010K08.4"

mRNA 13662..14499
 /gene="OSJNBa0010K08.4"
 /note="supported by full-length cDNA(s) : AK069347"
 CDS 13704..14132
 /gene="OSJNBa0010K08.4"
 /note="contains EST(s) : AU174440 (E70063), AU174441 (E70063)
 contains full-length cDNA(s) : AK069347"
 /codon_start=1
 /product="eukaryotic translation initiation factor 1A"
 /protein_id="BAD23331.1"
 /db_xref="GI:48716665"
 /translation="MPKNGKGGKRRKRRKRNDEADKRELVFKEQGEYAQAARMIGN
 GREACDIOGTGKRLCHIRGKMKKRVIAAGDYLVLQYDQDKADAVILKYNDDEARL
 LKYGGEIPIHVRINBGEVDEDDAADDYDIOFEDEDIDKI"
 15912..16720
 /gene="OSJNBa0010K08.5"
 /join(15912..16111,16384..16720)
 /note="start and end point are not identified"
 /join(15912..16111,16384..16720)
 /gene="OSJNBa0010K08.5"
 /note="predicted by FGENESH etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD23332.1"
 /db_xref="GI:48716665"
 /translation="MAGRGKVDVPLRLDLAPRSQGCGNNPGMAARRHDDGVA
 AMCGCGGAOKSPAYRRPWRSGRHVVOSRVLWPKLVKGLAGGTEBAMASFO
 GRCQGYAACVYKGRRGASAVGRPPAALAGAVSLMKKQASKVYGAFGESLAWF
 IGPATVPASQPLGRG"
 21207..21747
 /gene="OSJNBa0010K08.6"
 /join(21207..21303,21342..21586,21730..21747)
 /gene="OSJNBa0010K08.6"
 /note="start and end point are not identified"
 /join(21207..21303,21342..21586,21730..21747)
 /gene="OSJNBa0010K08.6"
 /note="predicted by GENSCAN etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD23333.1"
 /db_xref="GI:48716666"
 /translation="MSSSKSKAINTSTTLVAASREBRARRKIQEMTWIPLSSGY
 AVERWCGLVASPPSPAPQRAAGFVAKINDNSHSLPLGRKRGKREATLSLSLLA
 TLRHCRASSPSLLILA"
 complement(22619..23165)
 /gene="OSJNBa0010K08.7"
 /complement(22619..23165)
 /gene="OSJNBa0010K08.7"
 /note="probably inactive due to 5' exon missing in CDS
 probably inactive due to including frameshift(s) in CDS
 pseudogene, lipoxigenase L-2"
 /pseudo
 /join(23876..24965,28559..28921)
 /gene="OSJNBa0010K08.8-1"
 /join(23876..24965,28559..28921)
 /gene="OSJNBa0010K08.8-1"
 /note="supported by full-length cDNA(s) : AK106298"
 /join(24093..24965,28559..28666)
 /gene="OSJNBa0010K08.8-1"
 /note="contains EST(s) : C26289 (C12036), AU166406 (C12036)
 contains full-length cDNA(s) : AK106298"
 /codon_start=1
 /product="putative zinc finger protein"
 /protein_id="BAD23334.1"
 /db_xref="GI:48716667"
 /translation="MADGGGGBAGSGGAPVCSFYRKKPKRIKRPAPAGDDDE
 DSGAIAARAKAPSTSKLFSSADGSSPPRFYESSRTIQATSDSATATLLE
 TFFDRPARIREROLQAEESSLKNPASAGSGSGEYKGIHGYTVYKAGFRREH
 TVSEKAGSGHGPLRASAHIRLSARPDYDPCDKYKTEGCGYGSCKMHDRGDYK
 SMOIEKEWEBAEKARKRRIAMGGGSDVEAGDEDDDEBALPFCYCYCREFPDVPV
 YTKCHYFCEHCAKHSKNCPCVQNKRTLGI FNAAGIRKKMADKKQ"
 28390..28918

```

mRNA      /gene="OSJNBa0010K08.8-2"
<28390..>28918
/note="OSJNBa0010K08.8-2"
misc_feature /note="supported by full-length cDNA(s) : AK063049"
28390..28918
/gene="OSJNBa0010K08.8-2"
/note="contains full-length cDNA(s) : AK063049
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
complement(join(34424..34584,34767..34818,34903..35051,
36403..36567,36919..37030,37086..37128,37297..37370))
/gene="OSJNBa0010K08.9"
complement(join(34424..34584,34767..34818,34903..35051,
36403..36567,36919..37030,37086..37128,37297..37370))
/gene="OSJNBa0010K08.9"
/note="hypothetical ORF
predicted by GENSCAN
this category is not included in IRGSP standard"
gene      complement(37645..41835)
/gene="OSJNBa0010K08.10"
mRNA      complement(join(37645..38343,39677..39796,40494..40872,
41186..41222,41582..41831))
/gene="OSJNBa0010K08.10"
/note="supported by full-length cDNA(s) : AK100174"
complement(join(37750..38343,39677..39796,40494..40872,
41186..41222,41582..41835))
/gene="OSJNBa0010K08.10"
/note="supported by full-length cDNA(s) : AK071676"
complement(join(37750..38343,39677..39796,40494..40872,
41186..41222,41582..41834))
/gene="OSJNBa0010K08.10"
/note="supported by full-length cDNA(s) : AK104411"
complement(join(38143..38343,39677..39796,40494..40872,
41186..41222,41582..41624))
mRNA      /gene="OSJNBa0010K08.10"
41186..41222,41582..41834))
/gene="OSJNBa0010K08.10"
/note="supported by full-length cDNA(s) : AK104411"
complement(join(38143..38343,39677..39796,40494..40872,
41186..41222,41582..41624))
CDS       /gene="OSJNBa0010K08.10"
41186..41222,41582..41624))
Query Match 2.5%; Score 273.8; DB 8; Length 184760;
Best Local Similarity 65.9%; Pred. No. 2e-25;
Matches 446; Conservative 0; Mismatches 222; Indels 9; Gaps 3;

```

```

Db 17979 TTAAT---ATCTGACGACTAGATCGACAGCGGAGGAAGTGGGATAGGTG 179854
Oy 1558 ATTGCTACGAGGATATACGCAAGCTGTTGGCGAGCGCCCGCTTATCCCATGGGA 2017
Db 179855 ATCAGCGACGAGGGGTACGACACTTCGCTTCGGAGCAAGCCGTTGTCGATGGG 179914
Oy 2018 GTGTTGGGACATCACCCCTGTCGTGATCATAGAGGTCCTGTCACAGTCATGATAGTG 2077
Db 179915 GTGTTGGGACGAGCGGCGCCGCTGATAGCGTGGCGGCAATCCAGCGGTGATG 179974
Oy 2078 CTGATGAGCGGCTTGATGAGGTGATGACACCCAGAAAGATCTTACAGAAACT 2137
Db 179975 CCGGCTGCGCGCTCGGCTGATGCGCGCACCGATCCAGATCTCAGAAACAG 180034
Oy 2138 AAGTACTTAATCTCT 2154
Db 180035 AAGGTCTTAATTTGCT 180051

RESULT 11
AP004679 111173 bp DNA linear PLN 24-JUN-2004
LOCUS      Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6,
DEFINITION BAC clone:OSJNBa0021N09, complete sequence.
ACCESSION  AP004679
VERSION     AP004679.2 GI:49169756
KEYWORDS    HTG.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretoidae; Oryzeae; Oryza.
REFERENCE   1
  Sasaki,T., Matsumoto,T. and Yamamoto,K.
  Oryza sativa nipponbare (Ga3) genomic DNA, chromosome 6, BAC
  clone:OSJNBa0021N09
  Published only in Database (2002)
  2 (bases 1 to 111173)
JOURNAL    Sasaki,T., Matsumoto,T. and Yamamoto,K.
AUTHORS    Direct Submission
TITLE      Submitted (30-JUN-2002) Takuji Sasaki, National Institute of
JOURNAL    Agrobiological Sciences, Rice Genome Research Program; Kannonnai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:sasaki@nias.affrc.go.jp, URL:http://xgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jun 23, 2004 this sequence version replaced gi:18447938.
The orientation of the sequence is from -21m3 to M3rev of the BAC
clone.
FEATURES   location/Qualifiers
            1..111173
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="genomic DNA"
            /cultivar="Nipponbare"
            /db_xref="taxon:39947"
            /chromosome="6"
            /clone="OSJNBa0021N09"
ORIGIN
Query Match 1.9%; Score 213; DB 8; Length 111173;
Best Local Similarity 60.4%; Pred. No. 1.3e-17;
Matches 408; Conservative 0; Mismatches 255; Indels 12; Gaps 3;

```


QY 1655 GCCCGCTTCAACAGAGCTGAGAGTCCGGCATTTGCACCTCATCCCGCAAGAGGGGTGGAG 1714
 DB 41702 GCGGTGTTCCACAGAGATGAGAGTGGGTCTATATCTTTGTTCCAGAGAGAGATGGAG 41761
 QY 1715 ATGACATCGATCGCTGGAAATCCATGCGCCGCAAGAACACCCGCGCATGTCATATA 1774
 DB 41762 ATTATGTTGAAGCTGTTGAAGCTTTAGCATATGATATCTGTTGCAATGATGATACT 41821
 QY 1775 AACCCCAACACCGGTGGGAGGGTTACTCTTACGACCATCTGTCCAAGGT---TTCA 1831
 DB 41822 AACCCCAATACCTTTGTTGTAATGTGACCTTATGACATCTGTCCAGAGTGGTATG 41881
 QY 1832 CATCTTTCCTTCCTGATATGATATGATTCAGTTCAGTGCACCTGCTGAATCTTTTCCCA 1891
 DB 41882 ATGCTGTACAAATCTGGGTAAACCTTACTTCAATTTATTTTTCATTTTCAATGAG 41941
 QY 1892 ATGCAATACGATCGATGTTGCTCAATTAAGTCCGCGAGGTGGGAAAGGCTGGAAATA 1951
 DB 41942 CTATATGTCCTATC-----TTAATCTAGATTCAGATACAGAACGCAAGTGTGCTG 41995
 QY 1952 TTGGTATGCTGACGAGTATACGAGCAAGCTGTGTTGCGGAGGCGCGCTGATCCCA 2011
 DB 41996 TTATGCTATGCTGATGAAGTATATGCTACCTTTGTTATGAGACACTCTTTTGTGCA 42055
 QY 2012 ATGGAAGTGTGGGCAATGACCCCTGTGCTGTCCATAGGGTCTGTCCAGTCAATGG 2071
 DB 42056 ATGGGTGTGTTGAGAGACTGTTCCAGTACTCAATTTGGAGCTATATCAAGATGG 42115
 QY 2072 ATATGCTGTGATGCGCGCTGTGATGGGTAGCGGTGACAGCCCAAGAAATCTTACAG 2131
 DB 42116 GTTGTCTCTGCTGAGATTTGGATGATGATGCAATTTGTATCTTAAAGGCAATCTGAAA 42175
 QY 2132 GAACTAAGTACTT 2146
 DB 42176 GAAACAAAGTGATTT 42190

RESULT 12
 AP003542/c 147640 bp DNA linear HTG 21-MAR-2002
 LOCUS Oryza sativa (japonica cultivar-group) chromosome 6 clone F0589C03,
 DEFINITION *** SEQUENCING IN PROGRESS ***.
 ACCESSION AP003542
 VERSION AP003542.1 GI:13810548
 KEYWORDS HTG; HTGS PHASE2.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eubartoideae; Oryzaceae; Oryza.

REFERENCE
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Oryza sativa nipponbare (GR3) genomic DNA, chromosome 6, PAC
 clone: F0589C03
 JOURNAL Published Only in Database (2001)
 REFERENCES 2 (bases 1 to 147640)
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 JOURNAL Direct Submission
 Submitted (25-APR-2001) Takuji Sasaki, National Institute of
 Agrobiological Resources, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@bar.affrc.go.jp, URL: http://rjp.dna.affrc.go.jp/
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and

* the accession number will be preserved.
 FEATURES
 source
 1. 147640
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="6"
 /clone="F0589C03"

ORIGIN
 Query Match 1.9%; Score 213; DB 2; Length 147640;
 Best Local Similarity 60.4%; Pred. No. 1.3e-17;
 Matches 408; Conservative 0; Mismatches 255; Indels 12; Gaps 3;

QY 1475 AAACACAGCGCGGTGGAGAGACCTGTGACAGGGGTGCGTACATGCTATGCGCGAC 1534
 DB 20788 AAATCCAGAGCTATTTGAGAGTACTTCTGTATCTTCTTCAAGCTTTGACAGAT 20729
 QY 1535 GAGCTCTTCTCACCGCGCGGAGCCAGGCGATGAGGTCAATATCCCGTGTGCGC 1594
 DB 20728 GATATTTTCTTCACTGTGAGGTACCAAGCAATGAGATGTTATGCTGTTTTGGC 20669
 QY 1595 CAGACCGCGCGCCCAATTTGCTCTCCAGGCGCAGGCTTACCCAAATACAGGCGCGC 1654
 DB 20668 CA---ACGAGGTCCCAATATATGCTTCCAAAGCCCGGTACCCAAACATGAGCAGAT 20612
 QY 1655 GCCCGCTTCAACAGCGTGGAGGTCGGGCAATTCGACTCATCCCGCAAGAGGGGTGGAG 1714
 DB 20611 GCGGTGTTCCACAGATGAGATGAGAGTGGGTCTATGATCTTTTCCAGAGAGATGGAG 20552
 QY 1715 ATGACATGACATCGCTGGAATCATCGCGAGAACACCAACCGCATGTCATATA 1774
 DB 20551 ATTATGTTGAAGCTGTGTAAGCTTTAGCAGATGAGAAATCTGTGCAATGATGATACT 20492
 QY 1775 AACCCCAACACCGGTGGGAGGCTTCTCTAGACCATCTGTCCAAAGT---TTCA 1831
 DB 20491 AACCCCAATACCTTTGTGTATGTCATCTTATGACATGAGCATGTGTCCAAGGTTGTATG 20432
 QY 1832 CATCTTTCCTTCCTGATATGATGATTCAGTTCAGTGCACCTGCTGAATCTTTTCCCA 1891
 DB 20431 ATGCTGTACAAATCTGGGTAAACCTTACTTATTTATTTTCCAAATTTCAATGGAG 20372
 QY 1892 ATGCAATACGATCGATGTTGCTCAATTAAGTCCGAGGTGCGAAGAGCTGGAAATA 1951
 DB 20371 CTATATGTCCTATC-----TTAATCTAGATTCAGATACAGAACGCAAGTGTGCTG 20318
 QY 1952 TTGGTATGCTGACGAGTATACGCGCAGAGCTGTTCTGGGAGCGCGCTTCATCCCA 2011
 DB 20317 TTATGCTATGCTGATGAAGTATATGTCACCTTGTATGAGAGCACTCTTTTGTGCCA 20258
 QY 2012 ATGGAAGTGTGGGCAATGACCCCTGTGCTGTCCATAGGGTCTGTCCAGTCAATGG 2071
 DB 20257 ATGGGTGTGTTTGAAGAGACTGTTCCAGTACTCAATTTGGAGCTATATCAAGATGG 20198
 QY 2072 ATATGCTGTGATGCGCGCTGTGATGGGTAGCGGTGACAGCCCAAGAAATCTTACAG 2131
 DB 20197 GTTGTCTCTGCTGAGATTTGGATGATGATGCAATTTGTATCTTAAAGGCAATTCGAAA 20138
 QY 2132 GAACTAAGTACTT 2146
 DB 20137 GAAACAAAGTGATTT 20123

RESULT 13
 ATP2009 93695 bp DNA linear PLN 03-AUG-1999
 LOCUS Arabidopsis thaliana DNA chromosome 4, BAC clone F2009 (ESSA
 DEFINITION project).
 ACCESSION AL021749
 VERSION AL021749.1 GI:2842474
 KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1
REFERENCE Bevan, M., Rose, M., Hempel, S., Entian, K.-D., Hohleisel, J.,
Mewes, H.-W., Meyer, K.F.X., Lemcke, K. and Schueller, C.
AUTHORS Unpublished
JOURNAL 2 (bases 1 to 93695)
REFERENCE EU Arabidopsis sequencing project.
AUTHORS Direct Submission
TITLE Submitted (30-JUN-1999) MIPS, at the Max-Planck-Institut fuer
JOURNAL Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
TITLE schueller@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
JOURNAL Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.
FEATURES
source location/Qualifiers
1..93695
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
18..467
/gene="P2009.10"
complement(18..467)
/gene="P2009.10"
complement(18..467)
/gene="P2009.10"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA16872.1"
/db_xref="GI:2842475"
/db_xref="UniProt/TREMBL:O49442"
/translation="MTSYKPLRVISKRRTACFLVYVAILSSNMSINSIYMWY
ENHQTSSGLPAIYASVLGAVFGVLSMAALPIAIPAVIIVITISVVTMAIRGSRKR
VVEGRKVTKEIAGFVFLVLEKGNFVLLCALLVFVFNSSSS"
complement(18..467)
/gene="P2009.10"
/number=1
2196..2675
/gene="P2009.20"
complement(2196..2675)
/gene="P2009.20"
complement(2196..2675)
/gene="P2009.20"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA16873.2"
/db_xref="GI:5701784"
/db_xref="UniProt/TREMBL:O49443"
/translation="MAKKEOETMSYFLPLRVISKRRTWCLFVVVAVILSSRNS
INSTVWYGENHQTSSGLPAIYASVLGAVFGVLSMAALPIAIPAVIIVITISVVTI
APVYSRKRVIVIEGRKVTKEIAGVIFKVLKGNFVLLCAVIAIVFVFNSSSS"
complement(2196..2675)
/gene="P2009.20"
/number=1
4143..6188
/gene="P2009.40"
join(4143..4991,4987..5044,5096..6188)
/note="similarly to lectin receptor-like serine/threonine
kinase lecrk1, Arabidopsis thaliana, PIR2:568589
Contains Protein kinases signatures and profile, pos.
328-349 & pos. 444-456"
/codon_start=1
/product="receptor protein kinase like protein"
/protein_id="CA16875.1"
/db_xref="GI:2842478"

/db_xref="GOA:O49445"
/db_xref="UniProt/TREMBL:O49445"
/translation="WFSKVSITLLPSLASLIFRSTTGIEFTYNSNFTTTLLGNAT
VKSPPSITLTINQTTPIGRGLYPSRINASSSAPLPFAFSFLPSMAPFRLSGHG
FAFVFLPSESESAASSQHLGFNFNNNGDNSRIFAEPVFNQENNDINDNVG
DVNSLTVASSTAGVGGDQRFTEELKNGENQAVIENFGSAINTVMAASRKP
IRPLISIPLNTGLVLDMPFVGFATSGLOVSHRILMSNSNSFSIGDVLKSGF
IAGVSGVVLQRLGDVDEWETEVPHRVQKDYLEATKGSDEMNGVCGNSKVRG
VLEGRVAVRIMNSPRESVATSEFLAVSLSLRIRKNIYGLKMSKKGESITLI
YEYENSGVDRIFDENMLWBERMRVIRDLASGMVLYHEGMEKVLHRIKSNVL
LDKDMNAVGDGFLAKLQNTSKENVSTHVGTAGVYAPLKVTKGRASQTDVYSGV
FLVRYCCRRPIEGREGIVEMINGMEKDKVVDGLDERIKANGVFWVEYEMALRIG
LVCVHPDRVAPRKROYVQVLEQRLVEDGERISLERVKSXYLETGSGSRQHP
TFOVWNSYSNSNFGYTDILHR"
4143..4991
/gene="P2009.40"
/number=1
4942..4986
/gene="P2009.40"
/number=1
4987..5044
/gene="P2009.40"
/number=2
5045..5095
/gene="P2009.40"
/number=2
5096..6188
/gene="P2009.40"
/number=3
6559..11254
/gene="P2009.30"
complement(join(6559..6669,6878..7050,7305..7428,
7537..7647,7763..7915,10629..10733,10823..10875,
11074..11254))
/gene="P2009.30"
complement(join(6559..6669,6878..7050,7305..7428,
7537..7647,7763..7915,10629..10733,10823..10875,
11074..11254))
/note="similarity to ENOD20 gene (small copper-binding
protein), Medicago truncatula, PATCH: E258656"
/codon_start=1
/product="copper-binding protein-like"
/protein_id="CA16874.2"
/db_xref="GI:5701785"
/db_xref="GOA:O49444"
/db_xref="UniProt/TREMBL:O49444"
/translation="MGLVMPFDLYLMEVLMGLGFTISNGYKPYVGKDGCVPTPSED
YSHNSHNRFOVNDTLPTOVTFSLVATLTKVLKELSSPGSPGVSESDNSRVAP
GPATAPASAGIGISTSRLOASBPVSSPALLSGKEBCKIIPKROKQAVIK
SIKOSPKKNIYALVYGMRYEDALLIQLTVYKRAAQTVYRHTTCFVNTLKGKYTI
LQVTHAANATHNGLDPRLLVIVRETPPEEAEIARLKVNFPKKSREKQLVP
HKLIETSPINRRGTKANHRSELVLTIIIDVTGVNMKNRLDNLITNNNIYHNPS
GP"
complement(6559..6669)
/gene="P2009.30"
/number=1
complement(6670..6877)
/gene="P2009.30"
/number=1
complement(6878..7050)
/gene="P2009.30"
/number=2
complement(7051..7304)
/gene="P2009.30"
/number=2
complement(7305..7428)
/gene="P2009.30"
/number=3
complement(7429..7536)
/gene="P2009.30"
complement(7537..7647)

```

      /gene="F2009.30"
      /number=4
      /complement(17648..7762)
      /gene="F2009.30"
      /number=4
      /complement(17763..7915)
      /gene="F2009.30"
      /number=5
      /complement(17916..10628)
      /gene="F2009.30"
      /number=5
      /complement(10056..10127)
      /gene="F2009.30"
      /product="rRNA-Lys"
      /note="codon recognized: AAA"
      /complement(10629..10733)
      /gene="F2009.30"
      /number=6
      /complement(10734..10822)
      /gene="F2009.30"
      /number=6
      /complement(10823..10875)
      /gene="F2009.30"
      /number=7
      /complement(10876..11073)
      /gene="F2009.30"
      /number=7
      /complement(11074..11254)
      /gene="F2009.30"
      /number=8
      /note="ESTmatch (T04156_F)"
      /complement(12220..12639)
      /gene="F2009.50"
      /complement(13176..16811)
      /gene="F2009.50"
      /complement(13187,13379..13519,13534..13707,14167..14233,
      14308..14385,14487..14635,14753..14783,15222..15422,
      15885..16117,16178..16213,16388..16586,16629..16739,
      16804..16811)
      /gene="F2009.50"
      /codon_start=1
      /product="hypothetical protein"
      /protein_id="CA16876.2"
      /db_xref="GI:5701786"
      /translation="MESKILRGRKQBAEVPFSAMNITGVYRPHALLHFFVYFYS
      GLSTMATNVDYLSNTRVRSCTKPLNSKSDPKXKNSVVELVAAPTETI
      TGVHYVQIGTFSQVQESPRDLKRLKLSVNEKHCNIEIAYQVSRYASSENNED
      1.9%; Score 212.6; DB 8; Length 93695;
      Best Local Similarity 53.2%; Pred. No. 1.4e-17;
      Matches 577; Conservative 0; Mismatches 459; Indels 49; Gaps 4;

```

```

      33617 TACTTGTTCACCCCACTGTAAAGTCTGTAAACCAATCAAAACCTAAAGATATTTTCATGCA 33676
      7013 TTATATAAAACAAACAGCGCCCTAGACAGACACTTGTCAACAGCGCGCCCTAACAACTAT 7072
      33677 ATGTGTTTGTAAACAGCGCGGTGTGACGATTTACGTAACCGAGCTTGAAGAACAAAGTAA 33736
      7073 CGCCGACAGACGCTCTCTCAACCGCGCGGAACTGAGCGATGAAAGTCATTAATCCCG 7132
      33737 AGCTTAATATATGATATTCATTAACGTGTGGAATGCAACCAAGGATAGAAAGTATGTTCAAG 33796
      7133 TGTGTGCCCAAGACTGCGCGGCCAATATCTGTTTCCCGGCGAGCTATCCAAATTAAG 7192
      33797 CACTGGCTC---GACCAACGCTAACATCTTCTCCAAAGGCTATGTCCTCACTACG 33853
      7193 AGCGGAGAGCGGATTCACAAAGCTGAGAGTCCGCACTTCGACTATCCCGGACAGAG 7252
      33854 AGGCTGTGCGGCTACAGGTGAGACTGAGGTCCGCAAGTTGATCTCTCCAGAGAAAG 33913
      7253 GTTGGAGATTCGACATCGACTCGCTGGAATTCATCGCCGACAAACACCAACCGCATGG 7312
      33914 AATGGAGATGATATCTCCAGGATGAAAGCCATGGAATGAAACAGTTGCATGG 33973
      7313 TCATCATTAACCAACCAATCCGTGCGGACAGCTTATCTTACGACATCTGCGCAAG 7372
      33974 TTATCATTAACCTTAACCAACCCCTGTGAAATGTTTACTCTTACGATCATCTTAAGAAAG 34033
      7373 TTTTGATTCATGATCATCTCTGCTGTGATGATGACCGGCTGTGTAACATATGATATG 7432
      34034 TCCGTATTC-----ATATCTTTAATTAAGTTGACTCGTGAAGACTGAAA 34080
      7433 GATTGCGTTGTCTAATCGTGTCTGATGATGCTGTTGTTATCAGTTCGCGGAGGTGGC 7492
      34081 TATTGGTGTGATGA-----ATGAACCTTTATAGGTGGCGAGACGCG 34124
      7493 AAGGAAGCTGGAAATTTGTGTATCCGTGACGAGGTTTACGCAAACTGTTCTGGGACG 7552
      34125 TAAAGAGCTGGGAATATGATATCAACAGCGAAGTTTATGCAAAACATCTTCGAGAG 34184
      7553 CGCCCGTTTATCCGATGAGGCGCTTTGGGCAATTGCCCGGCTGTGCAATGGATC 7612
      34185 CAATCTTTGTCTCCATGAGGAGAGTTTCTTGTGATATCCCGTTATCACTTTGGGCGG 34244
      7613 TCTGTCAAGTCTGTGATGATGCTGCTGATGACGACTTGAATGGTGGCGGTGACACCC 7672
      34245 TATATGAAAGATGATGATGTTCTCGGTGGAATATGGCTGATCGCTTGAATGACCC 34304
      7673 CACAAAGATTTAGAGAAACTAAGTAGCTTAACTCCCTATCATTTCTTATATGCT 7732
      34305 CAGAGGAATCTCAAGTCCACAGGGGTCTGTCTTTTACTGATCATCATTAAC 34364
      7733 ACTGT 7737
      34365 TGTGT 34369

```

RESULT 14
 ATCHR168
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
 1 (bases 1 to 5241; 88721 to 138031)
 Lemnard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
 Mewes H.W., Lemcke K. and Meyer K.F.X.
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 5137 to 88829)

AUTHORS Pohl,T., Weizenegger,T., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL Unpublished
REFERENCE 3 (bases 137918 to 199749)
AUTHORS Rose,M., Hempel,S., Entian,K.-D., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL Unpublished
REFERENCE 4 (bases 1 to 199749)
AUTHORS EU Arabidopsis sequencing project.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemccke@mips.biochem.mpg.de, mayermips.biochem.mpg.de, projectCOORDINATOR: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7JY Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/chal/> this fragment has an overlap with ATCHR169 at the 3' end and an overlap with ATCHR169 at the 3' end.
FEATURES
source location/Qualifiers
1.199749
/organism="Arabidopsis thaliana"
/mol_type="Genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="4"
6341..8260
/gene="AT4g27900"
join(6341..6448,7253..7708,7926..8066,8180..8260)
/gene="AT4g27900"
/note="similarity to zinc-finger protein S12569, Oryza sativa, PIR3:JE0115
containe EST gb:T42849, T43671, A1995239.1, T04160, R30019, T22597, T22233"
/codon_start=1
/product="putative protein"
/protein_id="CAB79593.1"
/db_xref="GI:7269645"
/db_xref="UniProt/TREMBL:Q9SUE8"
/translation="MYAETSLAFRPMOSSSPEIHQFEDLFKSYKLSDEMNVLVEASEYDFGEISDLFRAPEPIIEEPLMADPLSQELVGLDGLSQDQILDKAFYDEQDLVKSMSSEPLSEVLDIKNISLVAKNDNVKTSVVSVPKPSVSGNLSMNVETKDAVTCGPDPDPVDYCGWRAPSESDITCTGTGNTGYSQSLDRITICTSEBDRREKLSRYRNKSKRPFGRKIKYACRKALDSQPRIRGRPAKTEEMQK"
6341..6448
/gene="AT4g27900"
/number=1
6449..7252
/gene="AT4g27900"
/number=1
7253..7708
/gene="AT4g27900"
/number=2
7709..7925
/gene="AT4g27900"
/number=2
7926..8066
/gene="AT4g27900"
/number=3
8067..8179
/gene="AT4g27900"
/number=3
8180..8260
/gene="AT4g27900"
/number=4
10177..15419
/gene="AT4g27910"
join(10177..10956,11042..11097,11169..11403,11496..11624,11885..11953,12055..12314,12474..12610,12836..13025,13575..13720,13825..13929,14174..14238,14331..14541,14835..14897,14950..15039,15129..15215,15303..15419)
/gene="AT4g27910"

/note="similarity to MLF protein, Gallus gallus, Av011003
containe EST gb:217631"
/codon_start=1
/product="putative protein"
/protein_id="CAB79593.1"
/db_xref="GI:7269645"
/db_xref="GOA:Q9SUE7"
/db_xref="UniProt/TREMBL:Q9SUE7"
/translation="MIIRKKKTKQIPSLERKLGNEBRKKRKNLGGGGYYPANLLGEIAGIVPGNGRNGFSASCTEVTKEPVEVESLSKRSDGIVRDSPPAATMLKSRITSGRQVLPESRNDSDLDWVRDSDSDCLEEEIECRNEKVSFEPPATMLKSKELDKSKYSALCKEERFHEOHNDARARVDEKLNKGTGFEPEFVSGLVWASGRN EEPFPAVYIDMTQAPELVLRSCIPDAVCVFPFGHSGNENRDYAMVRGIPFVDYVAROOPBELQCGCKRGNOMLBEAPLADQFETKLHNDHLAAGNTPDSDPMDIOETAVSNGEILNNAPRQASVSNITLILKKRPLACAGETYSIEMAKKMDLPGDOLCKPCSRVRCDGCKVRVIAHCDQISPHLKITVLAEDQLTLVQIQKRTQSRPLRNLFYCFNLGELLIAKCNLTGVFWRVAKGQVLPKRVIVCGVGVFEPRLHLVVCCKGSCGPKKKKALSEWRHSGSKSKMKTIVKYSKLALDEWKEKESCNMLQCPCFCDAFVGFSPFCOMRMSRSGQNDVOPADGLKITGTRLLSATNGALKPTDVTIWHVHTGCAFPQPCVCPASFEKKMEPAGLISITSTNPVKICVICOIHGSCQCCSKCTYTHAMCASRAGTMEHLCLEKNGQITTRVSCATYHRAHPNDVLIQPSGAFSASLIVNKKKGGSRILSLIRDEDAPEVNTICDPSARCVFPRKINSKGTTSPLAKNSLIMFERDNNVORTENDRVCFORSGHGGLPARNTIOGEWLVYRGGQVRGSIADLRARVRYVKDCYLFKISEVVVDATDKNIIARLNIHVSCLAMTFYVQEI"
10177..10956
/gene="AT4g27910"
/number=1
10957..11041
/gene="AT4g27910"
/number=1
11042..11097
/gene="AT4g27910"
/number=2
11098..11168
/gene="AT4g27910"
/number=2
11169..11403
/gene="AT4g27910"
/number=3
11404..11495
/gene="AT4g27910"
/number=3
11496..11624
/gene="AT4g27910"
/number=4
11625..11884
/gene="AT4g27910"
/number=4
11885..11953
/gene="AT4g27910"
/number=5
11954..12054
/gene="AT4g27910"
/number=5
12055..12314
/gene="AT4g27910"
/number=6
12315..12474
/gene="AT4g27910"
/number=6
12475..12610
/gene="AT4g27910"
/number=7
12611..12835
/gene="AT4g27910"
/number=7
12836..13025
/gene="AT4g27910"
/number=8
13026..13574
/gene="AT4g27910"
/number=8
13575..13720

intron	/gene="ATAg27910" 13721. .13824 /gene="ATAg27910" /number=9
exon	13825. .13929 /gene="ATAg27910" /number=10
intron	13930. .14173 /gene="ATAg27910" /number=10
exon	14174. .14238 /gene="ATAg27910" /number=11
intron	14239. .14330 /gene="ATAg27910" /number=11
exon	14331. .14541 /gene="ATAg27910" /number=12
intron	14542. .14834 /gene="ATAg27910" /number=12
exon	14835. .14897 /gene="ATAg27910" /number=13
intron	14898. .14949 /gene="ATAg27910" /number=13
exon	14950. .15039 /gene="ATAg27910" /number=14
intron	15040. .15128 /gene="ATAg27910" /number=14
exon	15129. .15215 /gene="ATAg27910" /number=15
intron	15216. .15302

Query Match Similarity	1.9%	Score 212.6	DB 8	length 199749
Best Local Similarity	53.2%	Pred. No. 14-17		
Matches	577	Conservative	0	Mismatches 459; Indels 49; Gaps 4;
Qy	6653	CGGGCGAAGACGCACCTCTGGCGACGACGGGGCGAAGAACGACATCCGGGCGATACGG	6712	
Db	171251	CGTTTCAGGGGAAAGGACAAACGACGACAAAGCCTCCAGCGTCAAGATGAGATCATCGTC	171310	
Qy	6713	TACAAAGATGAGGCGGACGGGTGAGAGAGACGGGCGCGGCGCGTGCCTGGCGCCAC	6772	
Db	171311	TACAAAGCTTTGATGATGATGCAACCTCGAATGTGAAAAAGCTCTTTTACCCTTGGCTCAC	171370	
Qy	6773	GGTGAACCCGTCCTGTTCCCGGCTTCGCGACGGGCGTCGAGGCGGAGACGCGCTATGCC	6832	
Db	171371	GGTAGACCTTTCTGTCTACCTCTGTTATCCGACACTCATCTCGTCGAGAAACGCGCGTGGTC	171430	
Qy	6833	GCCGGCTGGCGGACCGGCGCAATTCATCTGTACAGCCGGGGCGTGGCGCTCCCCCGCGCA	6892	
Db	171431	GACGTCCTTCGCTCGCGCAAGGATACCTTTAGGGCGCCGCGCGAATTTCTCCAGCC	171490	
Qy	6893	CGAAGTAACTTTACAGCTTACCCGTAAATGTATGCTGAGCATGATGCGCGGTTTAC	6952	
Db	171491	AGACAGTAAATATACATAATTTTCTCTACTTACTC-----TTAAT	171533	
Qy	6953	TTACGTGCCGCGCTGTTCTTCCCGGTGCGTTCAAAATTTTAACTTCTATAGTACC	7012	
Db	171534	TACTTTGTTCCACCCACCTGTAAGTCTGTAAACATTCAMAAACCTMAACGATATTTTCATGGA	171593	
Qy	7013	TTATAAACAACAACAGCGCGTACGAGACACTGTCAACAGGGCGTCCCTAACGATAT	7072	
Db	171594	ATGTTTGTGTATCAGGGCGGTGCGAGATTTAGTGAAACGAGACTTGGACACAAAGTAA	171653	
Qy	7073	CGGCGGACGAGCTTCTCTACCGCGGGGGAATCGAGGAGATGAAATCATATATCCCG	7132	

Db	171654	AGCCTAATGATGTATTTCATTACTGTCGGATGCAACAGGGATAGAAAGTAAGTCTTCAGT		171713	
QY	7133	TGCTGGCCCAGACTGCGCGCCCAACATACCTCTTCCCGGCAGGCTATCCAAATTACG		7192	
Db	171714	CACGTGGCTC---GACCAAAACGTAAACATCTCTTCCCAAGGCGCTAGTACCTCACTACG		171770	
QY	7193	AGGCGCGAGCGGCATTTCAACAAGCTGGAAGTCCGGCACTTCGACCTCATCCCGACAAAG		7252	
Db	171771	AGGCTCGGCGGTCTACAGTGAAGCTCGAGGTCCGAAGTTGATCTCTTCCAGAAAG		171830	
QY	7253	GGTGGAGATCGACATCGACTGCTGGAATCCATCGCGCAGCAAAACACACCGCGCATGG		7312	
Db	171831	AATGGGAATTTGATCTCCACGGCATCGAAAGCATTGGCAGATGAACACAGTTGCATGG		171890	
QY	7313	TCATCATAAACCCAAACAATCCGTGCGGACGCTTTACTCTCAACGCAATCTGGCCAAG		7372	
Db	171891	TTATCATTAACCCCTAACCAACCCCTGTGMAATGTTTACTCTTACGATCATCTTAAGAA		171950	
QY	7373	TTTTCGATCCATGCAATCCTCTGCGCTCGTTGATCGACCGGTCTGTTGAACATATGATATG		7432	
Db	171951	TCGCTATTC-----ATACCTTTTATTAAGCTTTTGA CTCGTGAGAGCTGAAA		171997	
QY	7433	GATTCGCTTGGCTAATCGTNGTCTGATGATGCTGTTGGTATCAAGTGCAGGAGGTGC		7492	
Db	171998	TATTGCTGCTATGA-----ATGAACCTTTATATAGGTGGCCGAGACGGC		172041	
QY	7493	AAGGAAGCTCGGAATATTTGGTGATCGCTGACGAGGTTTACGCCAAACCTGCTTGGGCAG		7552	
Db	172042	TAAAGAGCTGGGAATAATAGTATATCAACAGACGMACTTTATTCGCAAAACAATCTTCGGA		172101	
QY	7553	CGCCCGCTTATCCCGATGGGGGCTCTTTGGGCAACATGCGCCCGGCTTCCATTGGATC		7612	
Db	172102	CAAACTTTTGTCCCAATGGGGAGATTTCTTCGATTAATCCGGTTATCACTTTGGCGG		172161	
QY	7613	TCTGTCCAAGTCGTGATAGTGCCTGATATGCGACTGTGATGGGTGGCGGTATACAGACC		7672	
Db	172162	TATATCGAAAGGATGATGTTGTTCCGTGTTGAGAGATTTGGCTGATGCGCTTTGAATGACC		172221	
QY	7673	CACAAAGATTTTAGAAGAAACTAAGGTAGCTTTAGCTCCCTATCATTTCTTCATATGCT		7732	
Db	172222	CAGAGAGATTTCTCAAGTCCACACAGGGGTCTGTCTTCTTTTACTGATCATCATTAAC		172281	
QY	7733	ACTGT 7737		7792	
Db	172282	TEGT 172286		172341	

[illegible]

TITLE Direct Submission
JOURNAL Submitted (01-AUG-2004) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India
REFERENCE 3 (bases 1 to 147540)
AUTHORS Yadav,M., Pandit,A., Gaikwad,K., Srivastava,S.K., Singh,H., Singh,A., Mendiratta,S.D., Singh,P.K., Singhal,V., Dalal,V., Pat,A.K., Ghazi,I.A., Dixit,R., Batra,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2004) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India
REFERENCE 4 (bases 1 to 147540)
AUTHORS Dalal,V., Pandit,A., Ghazi,I.A., Yadav,M., Singh,P.K., Singh,A., Mendiratta,S.D., Singh,H., Srivastava,S.K., Singhal,V., Dixit,R., Batra,K., Gaikwad,K., Sharma,T.R., Mohapatra,T. and Singh,N.K.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2004) IIRGS, NRC on Plant Biotechnology, Indian Agricultural Research Institute, LBS Centre, New Delhi, Delhi 110012, India
COMMENT On Sep 29, 2004 this sequence version replaced gi:50897315.
FEATURES
Source
Location/Qualifiers
1..147540
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="11"
/clone="OSUNBa0072P10"
/clone_11b="HindIII"
ORIGIN
Query Match 1.9%; Score 209.6; DB 8; Length 147540;
Best Local Similarly 56.6%; Pred. No. 3.5e-17;
Matches 429; Conservative 0; Mismatches 324; Indels 5; Gaps 2;
QY 1457 ATACCAACAAACAAAAACACAGCGCGCTGGCAGACGCTGTGCGAGGGGTGCGG 1516
DB 52641 ATCTTAAACGAGATTGAAATTCAGAGCTTATCGCGGCTTATTCGAGACTTGCA 52582
QY 1517 TACATGCTATCGCGCAGAGCTTCTCTCACCGCGCGGAGCAGCGCATGAGTTC 1576
DB 52581 TATGAGCTATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 52522
QY 1577 ATATATCCGGTGTGCGCGCAGACCGCGCGCAGCATTTGCTCCCAAGCCAGGCTAC 1616
DB 52521 ATCTGCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 52465
QY 1637 CCAAACTACGAGCGCGCGCGCGCTTCAACAGCGCTGAGAGTCCGGCATTTGACCTCATC 1696
DB 52464 CTGTTCCAGAGGACCGCGCGCTTCAATGAGTGAAGTCACTTTGATCTTCTC 52405
QY 1697 CCGGACAGGGGTGGAGATGACATCGACTGGAATTCATCGCCGAGCAAGAACACC 1756
DB 52404 CCAAGAGTGGCTGGAGTGTGATCTTGATGAGTGGAGAACTTGCTGACAAACACG 52345
QY 1757 ACCGCTATGCTCATTAACCCCAACACCCGCGCGAGCGTTTACTCTTACGACCAT 1816
DB 52344 GTTGCAATGCTATTAATCAACAGAAATCCCTGTGCAATGTCTACATTTGAGACAT 52285
QY 1817 CTGTCAAGGTTTCAATCTCTTCTGCTGGAATGATGATGATGATGATGATGATGATG 1876
DB 52284 TTGGCCAGAGCACTAGTAAATTAATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 52225
QY 1877 GAATCTTTTTCGAATGCTAGTACTGATGTGTCTCAATTAGTGTGCGAGAGTGGCG 1936
DB 52224 GGAATTTTCAAGTGTGTGATCAAGATCAAGTGT--TCTGTAGTGTGCGAGACCGCG 52167
QY 1937 AAAAGCTCGGAATATGCTGATGCTGAGAGGTATACGCAAGCTGTCTTGGGACG 1996
DB 52166 AAAAGCTCGCATCTTTGCTATGCAAGATGAATGATGACATTTGACCTTTGGGACG 52107

QY 1997 GCCCGTTATCCCATGGAGAGTGTGGGACATCAACCCCTGCTGTCTCATAGGGTCT 2056
DB 52106 AATTAATTTGTGCGGAGTGGGGCTGTGGCTCACTAGCTTCACTTACCTTGGATCG 52047
QY 2057 CTGTCAAGTCATGATAGTATGCTGTGATGGCGGCTTGAATGGCTGATGACAGCC 2116
DB 52046 ATATCGAAGAGATGGGTGTGCTGTGATGGGACCTTGATGATGATGATGATGATGAT 51987
QY 2117 AGAAGATCTTACAGGAACTAAGTACTTAATCTATATCTTCTTTCAATGCTA 2176
DB 51986 AATGCGCTATTTCAAGACCAAGAGTGTGATTAATCTTCAAAACCTTTTCAACAA 51927
QY 2177 CTAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2214
DB 51926 GTATTTTGCATATCAACAAACAAATCAAGATTTTGG 51889

Search completed: October 15, 2005, 01:33:35
Job time : 44828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2005, 13:04:36 / Search time 98.0375 Seconds
(without alignments)

11862.046 Million cell updates/sec

Title: US-10-019-783-3

Perfect score: 24
Sequence: 1 tctggagatcagtcgacccaaca 24

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_atg: *
12: gb_ay: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

1	24	100.0	10966	6	E49893	E49893 Creation of
2	24	100.0	10966	6	BD012010	BD012010 Creation
3	24	100.0	10966	8	AB024006	AB024006 Hordeum v
4	19.4	80.8	2969	2	AC022300	AC022300 Mus muscu
5	19.4	80.8	104147	9	AC004703	AC004703 Homo sapi
6	19.4	80.8	152056	9	AC146691	AC146691 Macaca mu
7	19.4	80.8	171952	10	AC002324	AC002324 Mus muscu
8	19.4	80.8	172940	10	AC126544	AC126544 Homo sapi
9	19.4	80.8	192550	2	AC025584	AC025584 Mus muscu
10	19.4	80.8	193004	10	AL591177	AL591177 Mouse DNA
11	19.4	80.8	198388	2	AC073426	AC073426 Homo sapi
12	19.4	80.8	206578	2	AC126546	AC126546 Homo sapi
13	19.2	80.0	57662	2	AC107969	AC107969 Homo sapi
14	19.2	80.0	85378	2	AL356577	AL356577 Homo sapi
15	19.2	80.0	98360	9	HSDJ247C2	AL049713 Human DNA
16	19.2	80.0	135353	9	AC146086	AC146086 Pan trogl
17	19.2	80.0	145264	9	AC107939	AC107939 Homo sapi
18	19.2	80.0	154344	9	AC073308	AC073308 Homo sapi
19	19.2	80.0	168139	2	AC016936	AC016936 Homo sapi
20	19.2	80.0	185259	2	AC063953	AC063953 Homo sapi
21	19.2	80.0	187204	9	AC084082	AC084082 Homo sapi
22	19.2	80.0	188417	2	AC150805	AC150805 Callithrix
23	19.2	80.0	219730	2	AC110957	AC110957 Rattus no
24	19.2	80.0	225033	2	AC125338	AC125338 Rattus no
25	19.2	80.0	235082	2	AC094542	AC094542 Rattus no
26	18.8	78.3	150	6	AX753668	AX753668 Sequence
27	18.8	78.3	500	6	AX753680	AX753680 Sequence
28	18.8	78.3	1026	6	CO745253	CO745253 Sequence
29	18.8	78.3	1689	6	AX753658	AX753658 Sequence
30	18.8	78.3	1826	6	CO719286	CO719286 Sequence
31	18.8	78.3	1867	6	AX082301	AX082301 Sequence
32	18.8	78.3	1932	6	BC014126	BC014126 Homo sapi
33	18.8	78.3	2271	6	AX576068	AX576068 Sequence
34	18.8	78.3	2610	6	AX753655	AX753655 Sequence
35	18.8	78.3	2830	6	BD270045	BD270045 Sequence
36	18.8	78.3	3173	6	AX753654	AX753654 Sequence
37	18.8	78.3	3248	6	AX753657	AX753657 Sequence
38	18.8	78.3	3418	9	BC037539	BC037539 Homo sapi
39	18.8	78.3	5296	9	AK074084	AK074084 Homo sapi
40	18.8	78.3	91200	9	AB001152	AB001152 Homo sapi
41	18.8	78.3	147556	2	AC011007	AC011007 Homo sapi
42	18.8	78.3	154324	2	AC122552	AC122552 Mus muscu
43	18.8	78.3	173235	2	AC023756	AC023756 Homo sapi
44	18.8	78.3	204954	10	AC116151	AC116151 Mus muscu
45	18.4	76.7	174965	9	AC036222	AC036222 Homo sapi

ALIGNMENTS

RESULT 1	E49893	10966 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	E49893	Creation of iron-deficiency resisting rice plant.			
DEFINITION	E49893	Creation of iron-deficiency resisting rice plant.			
ACCESSION	E49893.1	GI:18629372			
VERSION	JP 2001017012-A/1.				
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified				
REFERENCE	1 (bases 1 to 10966)				
AUTHORS	Mori, S., Nakanishi, N., Takahashi, M. and Nishizawa, N.				
TITLE	Creation of iron-deficiency resisting rice plant				
JOURNAL	Patent: JP 2001017012-A 1 23-JAN-2001;				
COMMENT	SCIENCE & TECH AGENCY, SATOSHI MORI				
	OS Hordeum vulgare L. var. Igr1				
	PN JP 2001017012-A/1				
	PD 23-JAN-2001				
	PR 05-JUL-1999 JP 1999190318				
	PI SATOSHI MORI, NOBUHITO NAKANISHI, MICHIO TAKAHASHI, PI NAOKO				
	NISHIZAWA				
	PC A01H5/00, C12N5/10, C12N15/09//C12N9/10, C12N5/00, C12N15/00 CC				

FEATURES	source	1..10966	Location/Qualifiers
			/organism="Hordeum vulgare L. var. Igr1".
ORIGIN			
Query Match	100.0%; Score 24; DB 6; Length 10966;		
Best Local Similarity	100.0%; Pred. No. 0.055;		
Matches	24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Cy	1 TCTGGATCTCAGTCATCCACA 24		
Db	8475 TCTGGATCTCAGTCATCCACA 8498		

RESULT 2	BD012010	10966 bp	DNA	linear	PAT 02-AUG-2002
LOCUS	BD012010	Creation of iron-deficiency resisting rice plant.			
DEFINITION	BD012010	Creation of iron-deficiency resisting rice plant.			
ACCESSION	BD012010.1	GI:22092199			
VERSION	WO 0101762-A/1.				
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 10966)				
AUTHORS	Mori,S., Nakanishi,H., Takahashi,M. and Nishizawa,N.				
TITLE	Creation of iron-deficiency resisting rice plant				
JOURNAL	Patent: WO 0101762-A 1 11-JAN-2001; JAPAN SCIENCE AND TECHNOLOGY CORP, SATOSHI MORI, HIROMI NAKANISHI, MICHIO TAKAHASHI, NAOKO NISHIZAWA				
COMMENT	OS Hordeum vulgare L. var. Igr1 PN WO 0101762-A/1 PD 11-JAN-2001 PF 04-JUL-2000 WO 2000JP004425 PR 05-JUL-1999 JP 99P 190318 PI SATOSHI MORI, HIROMI NAKANISHI, MICHIO TAKAHASHI, PI NAOKO NISHIZAWA PC A01H5/00, C12N5/14, C12N15/52 CC				
FEATURES	FH Key	Location/Qualifiers.			
source		1..10966 Location/Qualifiers			

ORIGIN			
Query Match	100.0%; Score 24; DB 6; Length 10966;		
Best Local Similarity	100.0%; Pred. No. 0.055;		
Matches	24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Cy	1 TCTGGATCTCAGTCATCCACA 24		
Db	8475 TCTGGATCTCAGTCATCCACA 8498		
RESULT 3	AB024006	10966 bp	DNA
LOCUS	AB024006	linear	
DEFINITION	AB024006	Hordeum vulgare naat-B and naat-A genes for nicotianamine	
ACCESSION	AB024006	complete cds.	
VERSION	AB024006.1	GI:6469088	
KEYWORDS	nicotianamine aminotransferase.		
SOURCE	Hordeum vulgare subsp. vulgare		
ORGANISM	Hordeum vulgare subsp. vulgare		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		


```

repeat_region /rpt_family="AT_r1ch"
25785..26070
/rpt_family="AluSx"
repeat_region complement(26071..26107)
/rpt_family="AT_r1ch"
repeat_region complement(26186..26477)
/rpt_family="AluT0"
repeat_region complement(26594..26757)
/rpt_family="MIR"
repeat_region 26902..27202
/rpt_family="AluSg"
repeat_region complement(28990..29296)
/rpt_family="AluT0"
repeat_region 29742..29896
/rpt_family="MIR"
repeat_region complement(30043..30592)
/rpt_family="MLT1F"
repeat_region complement(31090..31367)
/rpt_family="AluSx"
repeat_region 32074..32199
/rpt_family="MIR"
repeat_region 32206..32521
/rpt_family="AluSp"
repeat_region 32572..32757
/rpt_family="MIR"
repeat_region complement(32761..32838)
/rpt_family="MER2"
repeat_region complement(33407..33553)
/rpt_family="MER5B"
repeat_region complement(33575..33756)
/rpt_family="MER5A"
repeat_region complement(33771..33973)
/rpt_family="MER5B"
repeat_region 34187..34271
/rpt_family="MIR"
repeat_region 34988..35288
/rpt_family="AluSx"
repeat_region 35412..35477
/rpt_family="MER5A"
repeat_region 35620..35692
/rpt_family="L1PA2"
repeat_region 36302..36575
/rpt_family="L1MB5"
repeat_region 36577..36686
/rpt_family="MIR"
repeat_region complement(36933..37007)
/rpt_family="(TGA)A)n"
repeat_region 37875..37999
/rpt_family="FLAM_A"
repeat_region 38476..38556
/rpt_family="PTR5"
repeat_region 38557..38849
/rpt_family="AluY"
repeat_region 38850..39247
/rpt_family="L1R12"
repeat_region complement(39625..39748)
/rpt_family="MIR"
repeat_region 40650..40674
/rpt_family="(CAG)n"
repeat_region 42135..42188
/rpt_family="(GGA)n"
repeat_region complement(43811..43983)
/rpt_family="MIR"
repeat_region 44313..44493
/rpt_family="MER5B"
repeat_region complement(44546..44620)
/rpt_family="MER5B"
repeat_region complement(45397..45480)
/rpt_family="L1MB8"
repeat_region 45317..45584
/rpt_family="MLT1F"

```

Query Match

80.8%; Score 19.4; DB 9; Length 104147;

Best Local Similarity 95.2%; Pred. No. 27;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CTGGATCTCAGTCATCA 22
Db 45121 CTGGATCTCAGTCATCA 45101

RESULT 6

AC146691

LOCUS AC146691 152056 bp DNA linear PRI 03-FEB-2004
DEFINITION Macaca mulatta clone ch250-214a10, complete sequence.
AC146691
VERSION AC146691.5 GI:41684679

KEYWORDS

SOURCE HTG.
ORGANISM Macaca mulatta (rhesus monkey)

Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.

REFERENCE

1 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Rhesus macaque BAC clone ch250-214a10
Unpublished

REFERENCE

2 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.

REFERENCE

3 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (17-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

4 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (15-JAN-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

5 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (27-JAN-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE

6 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (03-FEB-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE

7 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (03-FEB-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE

8 (bases 1 to 152056)
Qin,B., Lyons,L. and Roe,B.A.
Direct Submission
Submitted (03-FEB-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT

On Feb 3, 2004 this sequence version replaced gi:41349941.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES
source location/Qualifiers
1..152056
/organism="Macaca mulatta"
/mol_type="genomic DNA"
/db_xref="taxon:9544"
/clone="ch250-214a10"
/clone_1b="CHORI-250 Rhesus macaque BAC library"

ORIGIN

Query Match

Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GGGATCTCAGTCATCA 24
Db 45828 GGGATCTCAGTCATCA 45848

RESULT 7

AC002324
 LOCUS 171952 bp DNA linear ROD 30-MAY-1998
 DEFINITION Mus musculus chromosome 11, clone 475_H_14, complete sequence.
 AC002324
 VERSION AC002324.1 GI:3169207
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
 REFERENCE Birren, B., Fasmann, K., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Mus musculus chromosome 11, clone 475_H_14
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 171952)
 AUTHORS Hawkins, T.L., Reeve, M.P., Christofferson, A., Birren, B.W., Fasmann, K.H., Lander, E.S., McKernan, K., Munro, C., Richardson, P., Barna, N., Chang, A., Cooke, P., Daly, M.J., Devon, K., Dewar, K., Forrest, C., Gage, D., Geraghty, K., Guitau, G., Hagos, B., Huang, J., Jacotot, L., Lane, M., Lee, K., Mackenzie, J., Margulis, N., McDermott, J., Molla, M., Moloney, N., Morrow, J., Nachman, A., Naylor, J., O'Connor, T., Olootu, A., Peterson, K., Rollins, G., Spencer, J., Stiliwell, U., Stone, C., Strickland, C., Sydney, K., Traish, A., Wilmer, F., Zemseva, I. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUL-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 171952)
 AUTHORS Birren, B., Fasmann, K., Linton, L., Nusbaum, C., Lander, E., Allen, N., Baker, J., Baldwin, J., Barna, N., Beckert, R., Ben, U., Boatn, C., Boutwell, C., Brown, A., Castle, A., Cerny, J., Cooke, P., Depayre, E., Devon, K., Dewar, K., Donelan, L., Etemad, S., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S., Gershengorn, S., Geraghty, K., Gilmarin, T., Grant, G., Hagos, B., Harris, K., Horton, L., Howland, J.C., Hul, L., Jacotot, L., Kam, L., MacDonald, P., Margulis, N., McKernan, P., McGuire, A., McKernan, K., Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nachman, A., Naylor, J., O'Connor, T., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Rossello, R., Roy, A., Shyam, R., Stange-Thomann, N., Stiliwell, U., Stojanovic, N., Stone, C., Strickland, C., Subramanian, A., Törnqvist, M., Vassiliev, H., Vo, A., Wegner, A., Wang, B., Wheeler, J., Wu, Y., Ye, W., Zhao, J. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAY-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On May 30, 1998 this sequence version replaced gi:3142319. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>.
 FEATURES
 source Location/Qualifiers
 1. 171952
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /map="11"
 /clone="475_H_14"
 /clone_lib="Research Genetics/Cal Tech CITB-HSP-C (plates 195-384)"
 repeat_region 371..456
 /rpt_family="PB1D10"
 complement(1712..827)
 repeat_region /rpt_family="PB1D10"
 complement(990..1029)
 repeat_region /rpt_family="PB1D9"
 complement(1309..1451)
 repeat_region /rpt_family="B1_MM"
 complement(1882..2071)
 repeat_region /rpt_family="URR1B"
 2523..2563
 /rpt_family="B3A"
 repeat_region 2536..2717

repeat_region /rpt_family="B3"
 complement(4356..4383)
 repeat_region /rpt_family="POLY_A"
 5275..5494
 repeat_region /rpt_family="B3"
 5501..5674
 repeat_region /rpt_family="B3"
 6287..6402
 repeat_region /rpt_family="PB1D10"
 complement(7238..7371)
 repeat_region /rpt_family="L2"
 7659..7737
 repeat_region /rpt_family="B1-F"
 complement(8284..8429)
 repeat_region /rpt_family="B1_MM"
 8576..8652
 repeat_region /rpt_family="L2"
 complement(11209..11411)
 repeat_region /rpt_family="B3"
 complement(11436..11605)
 repeat_region /rpt_family="B3"
 complement(11634..11836)
 repeat_region /rpt_family="B2"
 complement(11861..12013)
 repeat_region /rpt_family="B2"
 complement(12262..12303)
 repeat_region /rpt_family="PB1D9"
 12306..12347
 repeat_region /rpt_family="L2"
 complement(12573..12596)
 repeat_region /rpt_family="(CCAA)n"
 complement(12639..12762)
 repeat_region /rpt_family="B4"
 complement(12807..12872)
 repeat_region /rpt_family="PB1D7"
 12920..13046
 repeat_region /rpt_family="B1-F"
 complement(13101..13294)
 repeat_region /rpt_family="B2"
 13310..13532
 repeat_region /rpt_family="MER2"
 complement(13534..13650)
 repeat_region /rpt_family="B1_MM"
 complement(13597..13670)
 repeat_region /rpt_family="B4A"
 complement(13684..13829)
 repeat_region /rpt_family="B1-F"
 complement(13804..13837)
 repeat_region /rpt_family="MT2B"
 13897..13922
 repeat_region /rpt_family="MT2B"
 complement(13899..14050)
 note="Single-stranded terminator coverage."
 13923..14070
 repeat_region /rpt_family="B1_MM"
 complement(14207..14351)
 repeat_region /rpt_family="B1_MM"
 complement(14958..15399)
 repeat_region /rpt_family="MTC"
 16021..16143
 repeat_region /rpt_family="PB1D10"
 16153..16236
 repeat_region /rpt_family="PB1D7"
 complement(17105..17224)
 repeat_region /rpt_family="(CA)n"
 complement(17227..17290)
 repeat_region /rpt_family="(CA)n"
 complement(18858..18900)
 repeat_region /rpt_family="AT_rich"
 complement(18901..18907)
 repeat_region /rpt_family="B1-F"
 complement(18908..18966)
 repeat_region /rpt_family="B4"

```

repeat_region      complement(18967, .19010)
                    /rpt_family="B1-F"
repeat_region      complement(19543, .19690)
                    /rpt_family="RSINE1"
repeat_region      complement(19713, .19815)
                    /rpt_family="PB1D10"
repeat_region      19950, .20032
                    /rpt_family="PB1D9"
repeat_region      21057, .21132
                    /rpt_family="PB1D7"
repeat_region      21151, .21201
                    /rpt_family="ID4"
repeat_region      21438, .21541
                    /rpt_family="B1-F"
repeat_region      complement(24457, .24557)
                    /rpt_family="PB1D9"
repeat_region      25092, .25226
                    /rpt_family="B1-MM"
repeat_region      complement(25295, .25403)
                    /rpt_family="PB1D7"
repeat_region      complement(25668, .25852)
                    /rpt_family="RSINE1"
repeat_region      complement(25985, .26117)
                    /rpt_family="B1-MM"
repeat_region      26126, .26176
                    /rpt_family="B3A"
repeat_region      complement(26292, .26359)
                    /rpt_family="ID4"
repeat_region      complement(28326, .28816)
                    /rpt_family="L1-MM"
repeat_region      29393, .29599
                    /rpt_family="B2"
repeat_region      29603, .29823
                    /rpt_family="B3"
repeat_region      29996, .30193
                    /rpt_family="B3"
repeat_region      complement(30412, .30563)
                    /rpt_family="B1-MM"
repeat_region      complement(31242, .31293)
                    /rpt_family="PB1D7"
repeat_region      31515, .31647
                    /rpt_family="B1-MM"
repeat_region      32490, .32579
                    /rpt_family="PB1D10"
repeat_region      33376, .33409
                    /note="Single-stranded terminator coverage."
repeat_region      33580, .33671
                    /rpt_family="GC rich"
repeat_region      complement(35248, .35318)
                    /rpt_family="(TA)n"
repeat_region      35664, .35809
                    /rpt_family="RSINE1"
repeat_region      36195, .36304
                    /rpt_family="PB1D9"
repeat_region      complement(36426, .36535)
                    /rpt_family="PB1D9"

Query Match      80.84; Score 19.4; DB 10; Length 171952;
Best Local Similarity 95.24; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy      4 GGGATCTCAGTGCATCCACA 24
Db      73085 GGGATCTCAGTGCATCCACA 73105

RESULT 8
AC126544
LOCUS      Homo sapiens chromosome 17, clone RP11-707023, complete sequence.
DEFINITION
AC126544
ACCESSION      AC126544.5 GI:24111035
VERSION
KEYWORDS      Homo sapiens (human)
SOURCE

```

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 172940)

Blumen, B., Nussbaum, C., and Lander, E.

Homo sapiens chromosome 17, clone RP11-707023

Unpublished

REFERENCE

2 (bases 1 to 172940)

Blumen, B., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,

Barnes, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,

Camata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,

Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,

Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,

McCarthy, M., Meidrim, J., Menus, L., Milova, T., Mlenga, V.,

Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,

Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,

Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,

Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,

Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,

Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (06-JUL-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 172940)

Blumen, B., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,

Barnes, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,

Camata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,

Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,

Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,

McCarthy, M., Meidrim, J., Menus, L., Milova, T., Mlenga, V.,

Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,

Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,

Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,

Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,

Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,

Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (10-OCT-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 172940)

Blumen, B., Nussbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,

Barnes, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,

Camata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,

Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,

Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,

Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,

McCarthy, M., Meidrim, J., Menus, L., Milova, T., Mlenga, V.,

Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,

Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,

Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,

Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,

Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,

Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (18-OCT-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

TITLE

JOURNAL

JOURNAL

COMMENT

On Oct 18, 2002 this sequence version replaced gi:23683037.
All repeats were identified using RepeatMasker:
Smith, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: I27758

Center clone name: 707_O_23

Only the first 172,9 kilobases of this clone are being submitted.
The remainder overlaps accession number AC106030 [WIGR project
I24005].

FEATURES

source

1. .172940
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-707023"
/clone_1ib="RPC1-11 Human Male BAC"
complement(75..230)
/rpt_family="MIR"
repeat_region
1065..1341
/rpt_family="AluSx"
repeat_region
1344..1625
/rpt_family="AluSx"
repeat_region
1505..1520
/note="<30 qual SNGL region"
unsure
1531..1536
/note="<30 qual SNGL region"
unsure
1545..1554
/note="<30 qual SNGL region"
unsure
1558..1563
/note="<30 qual SNGL region"
unsure
1582..1589
/note="<30 qual SNGL region"
repeat_region
1849..2142
/rpt_family="AluSg"
repeat_region
2198..2493
/rpt_family="AluO"
complement(2556..2597)
repeat_region
/rpt_family="5S"
2719..3001
/rpt_family="AluSc"
repeat_region
3002..3075
/rpt_family="TTAA)n"
repeat_region
4446..4600
/rpt_family="L2"
repeat_region
complement(4691..4987)
/rpt_family="AluY"
5177..5197
/rpt_family="(CAAAA)n"
repeat_region
5318..5610
/rpt_family="AluSg"
repeat_region
6449..6732
/rpt_family="AluSx"
repeat_region
6745..6859
/rpt_family="FLAM_A"
complement(6872..7027)
/rpt_family="AluSg"
repeat_region
complement(7028..7092)
/rpt_family="MERSA"
7432..7565
/rpt_family="MERSA"
repeat_region
8173..8214
/rpt_family="L2"
repeat_region
complement(8589..8881)

repeat_region /rpt_family="AluSx"
complement(8882..8997)
/rpt_family="MIR"
repeat_region complement(9021..9375)
/rpt_family="L1MDa"
complement(9594..9668)
/rpt_family="L2"
repeat_region 9672..9956
/rpt_family="AluSc"
complement(10134..10281)
/rpt_family="(TA)n"
repeat_region 10282..10310
/rpt_family="(CA)n"
10311..10704
/rpt_family="L1ME1"
repeat_region 10707..10823
/rpt_family="FLAM_C"
10948..10965
/rpt_family="L1ME1"
complement(10966..11261)
/rpt_family="AluSx"
complement(11270..11570)
/rpt_family="AluSx"
repeat_region 11571..12181
/rpt_family="L1ME1"
12205..12506
/rpt_family="AluSx"
repeat_region 12516..12646
/rpt_family="FLAM_C"
12653..13174
/rpt_family="L1MA4"
repeat_region 13312..13506
/rpt_family="MERS8B"
13515..13735
/rpt_family="L1ME1"
repeat_region 13739..14044
/rpt_family="AluDb"
14047..14340
/rpt_family="AluY"
repeat_region 14366..14419
/rpt_family="L1ME1"
14420..14718
/rpt_family="AluO"
repeat_region 14719..14884
/rpt_family="L1ME1"
repeat_region 14885..15188
/rpt_family="AluY"
repeat_region 15189..15551

Query Match 80.8%; Score 19.4; DB 9; Length 172940;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTGGATCTCAGTCATCCAA 22

Db 109700 CTGGATCTCAGTCATCCAA 109720

RESULT 9

AC025584/c 192550 bp DNA linear HTG 06-SEP-2000
LOCUS Mus musculus clone RP23-263C10, WORKING DRAFT SEQUENCE, 20
DEFINITION unordered pieces.

ACCESSION AC025584

VERSION AC025584.3 GI:9972308

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 192550)
McCombie,W.R., Baker,J.P., Bahret,A., Bal,H., Dedhia,N.N., de la

Bastide,M., Huang,E.N., King,L., Kirchoff,K.A., Miller,B.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Naclimento, L.U., O'Shaughnessy, A.L., Preston, R.R., Rodriguez, M.A.,
Shar, R.S., Shakhhar, M., Spiegel, L.A., Toth, K. and Vill, M.D.
Mouse Genomic Sequence
Unpublished
2 (bases 1 to 192550)
McCombie, W.R.
Direct Submission
Submitted (11-MAR-2000) Lita Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
On Sep 6, 2000 this sequence version replaced gi:8439865.

----- Genome Center
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory

Center code: GSHL

Web site: <http://www.cshl.org/genseq>

Contact: mccombie@cshl.org

----- Project Information

Center project name: RP23-263C10

Center clone name: RP23-263C10

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 42099: contig of 42099 bp in length
42100 42499: gap of unknown length
42500 59653: contig of 17154 bp in length
59654 60053: gap of unknown length
60054 76545: contig of 16492 bp in length
76546 76945: gap of unknown length
76946 90821: contig of 13876 bp in length
90822 91221: gap of unknown length
91222 105073: contig of 13852 bp in length
105074 105473: gap of unknown length
105474 117737: contig of 12264 bp in length
117738 118137: gap of unknown length
118138 129384: contig of 10847 bp in length
129385 129389: gap of unknown length
129389 138289: contig of 8905 bp in length
138290 138689: gap of unknown length
138690 147015: contig of 8326 bp in length
147016 147415: gap of unknown length
147416 153692: contig of 6277 bp in length
153693 154092: gap of unknown length
154093 159704: contig of 5612 bp in length
159705 160104: gap of unknown length
160105 165289: contig of 5185 bp in length
165290 165689: gap of unknown length
165690 170332: contig of 4643 bp in length
170333 170732: gap of unknown length
170733 175202: contig of 4470 bp in length
175203 175602: gap of unknown length
175603 178792: contig of 3180 bp in length
178793 179192: gap of unknown length
179193 182135: contig of 2943 bp in length
182136 182535: gap of unknown length
182536 185383: contig of 2848 bp in length
185384 185783: gap of unknown length
185784 188478: contig of 2635 bp in length
188479 188878: gap of unknown length
188879 191019: contig of 2141 bp in length
191020 191418: gap of unknown length
191419 192550: contig of 1132 bp in length.

FEATURES
SOURCE

1. 192550
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="caxon:10090"

ORIGIN

/clone="RP23-263C10"

Query Match 80.8%; Score 19.4; DB 2; Length 192550;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

4 GGGATCTCAGTGCATCCACAA 24
|||||
Db 191770 GGGATCTCAGATCCACAA 191750

RESULT 10

AL591177/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:1630497.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Bm, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/Wormpep RP23-399H5 is
from the RPCI-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pPAC3.6.

FEATURES

source

1. 193004
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="caxon:10090"
/chromosome="11"
/clone="RP23-399H5"
/clone_11b="RPCI-23"

ORIGIN

Query Match 80.8%; Score 19.4; DB 10; Length 193004;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

4 GGGATCTCAGTGCATCCACAA 24
|||||

Db 11904 GGGATCTCAGATCCACAA 11884

RESULT 11


```

/note="assembly_name:Contig38
clone_end:SP6
vector_side:right"
misc_feature 146329..164594
/note="assembly_name:Contig39"
misc_feature 164695..196388
/note="assembly_name:Contig40"

ORIGIN
Query Match 80.8%; Score 19.4; DB 2; Length 196388;
Best Local Similarity 95.2%; Fred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 CTGGAGTCAGTCAGTCACCA 22
Db 183265 CTGGAGTCAGTCAGTCACCA 183245

RESULT 12
AC126546/c 206578 bp DNA 1linear HTG 03-OCT-2002
LOCUS Homo sapiens chromosome 15 clone RP11-733021 map 15, WORKING DRAFT
DEFINITION AC126546
SEQUENCE AC126546
AC126546.2 GI:23477883
VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 206578)
AUTHORS Birren,B., Nuebaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 15, clone RP11-733021
REFERENCE 2 (bases 1 to 206578)
AUTHORS Birren,B., Nuebaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
Barra,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camara,A.J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Gardyna,S., Gord,S., Grahm,L., Grand-Pierre,N., Hagos,B.,
Horcun,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karcas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Melgrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Plerre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 206578)
AUTHORS Birren,B., Nuebaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
Barra,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camara,A.J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Gardyna,S., Gord,S., Grahm,L., Grand-Pierre,N., Hagos,B.,
Horcun,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karcas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Melgrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Plerre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,

```

TITLE JOURNAL COMMENT

```

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (03-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 3, 2002 this sequence version replaced GI:21700691.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27774
Center clone name: 733-O_21
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 20135 bases at least Q40
Consensus quality: 203359 bases at least Q30
Consensus quality: 204367 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 205078; sum-of-ctigs
Quality coverage: 15.0 in Q20 bases; agarose-fp
Quality coverage: 13.7 in Q20 bases; sum-of-ctigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1796: contig of 1796 bp in length
1896: gap of 100 bp in length
1897 2574: contig of 678 bp in length
2575 2674: gap of 100 bp
2675 3678: contig of 1004 bp in length
3679 3778: gap of 100 bp
3779 4472: contig of 694 bp in length
4473 4572: gap of 100 bp
4573 5587: contig of 1015 bp in length
5588 5687: gap of 100 bp
5688 6417: contig of 730 bp in length
6418 6517: gap of 100 bp
6518 7278: contig of 761 bp in length
7279 7378: gap of 100 bp
7379 8745: contig of 1367 bp in length
8746 8845: gap of 100 bp
8846 9956: contig of 1111 bp in length
9957 10056: gap of 100 bp
10057 11931: contig of 3875 bp in length
11932 14031: gap of 100 bp
14032 22254: contig of 8223 bp in length
22255 22354: gap of 100 bp
22355 29263: contig of 6909 bp in length
29264 29363: gap of 100 bp
29364 64453: contig of 35090 bp in length
64454 64553: gap of 100 bp
64554 91996: contig of 27443 bp in length
91997 92096: gap of 100 bp
92097 124248: contig of 32152 bp in length
124249 124348: gap of 100 bp
124349 206578: contig of 82230 bp in length.
-----
FEATURES
source
1..206578
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

/misc_feature /chromosome="15"
/misc_feature /map="15"
/misc_feature /clone="RP11-733021"
/misc_feature /clone_id="RC1-11 Human Male BAC"
/misc_feature 1.1796
/misc_feature /note="assembly_fragment"
/misc_feature clone_end:SP6
/misc_feature vector_side:left"
/misc_feature 1897..2574
/misc_feature /note="assembly_fragment"
/misc_feature 2675..3678
/misc_feature /note="assembly_fragment"
/misc_feature 3779..4472
/misc_feature /note="assembly_fragment"
/misc_feature 4573..5587
/misc_feature /note="assembly_fragment"
/misc_feature 5688..6417
/misc_feature /note="assembly_fragment"
/misc_feature 6518..7278
/misc_feature /note="assembly_fragment"
/misc_feature 7379..8745
/misc_feature /note="assembly_fragment"
/misc_feature 8846..9956
/misc_feature /note="assembly_fragment"
/misc_feature 10057..13931
/misc_feature /note="assembly_fragment"
/misc_feature 14032..22254
/misc_feature /note="assembly_fragment"
/misc_feature 22355..29263
/misc_feature /note="assembly_fragment"
/misc_feature 29364..64453
/misc_feature /note="assembly_fragment"
/misc_feature 64554..91996
/misc_feature /note="assembly_fragment"
/misc_feature 92097..124248
/misc_feature /note="assembly_fragment"
/misc_feature 124349..206578
/misc_feature /note="assembly_fragment"

```

ORIGIN

```

Query Match      80.8%; Score 19.4; DB 2; Length 206578;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy      2 CTGGGATCTCAGTCATCCAA 22
        |||||
Db      156371 CTGGGATCTCAGTCATCCAA 156351

```

RESULT 13

```

AC107969      57662 bp      DNA      linear      HTG 24-JAN-2002
LOCUS      Homo sapiens chromosome 11 clone CTD-2028015 map 11, LOW-PASS
DEFINITION      SEQUENCE SAMPLING.
ACCESSION      AC107969
VERSION      AC107969.1 GI:18308697
KEYWORDS      HTG; HTGS PHASEO.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      Birtten,B., Linton,L., Nusbaum,C. and Lander,E.

```

```

JOURNAL      1 (bases 1 to 57662)
REFERENCE      Homo sapiens chromosome 11, clone CTD-2028015
AUTHORS      Unpublished
                2 (bases 1 to 57662)
                Birtten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
                Anderson,S., Barina,N., Basilien,V., Bogunlavkiy,L., Boukhalter,B.,
                Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
                Chopel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
                Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
                Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,N.,
                Glnde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

```

TITLE
JOURNAL
COMMENT

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Margis,N., Matthews,C., McCarthy,M., McSwan,P., McKernan,K., Meldrum,J., Menus,L., Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phukhankang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Ries,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stranas,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: 124518

Center clone name: 2028_O15

```

***** NOTE: This record contains 73 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
718 717: contig of 717 bp in length
718 817: gap of 100 bp
818 1525: contig of 708 bp in length
1526 1625: gap of 100 bp
1626 2265: contig of 640 bp in length
2266 2365: gap of 100 bp
2366 3050: contig of 685 bp in length
3051 3150: gap of 100 bp
3151 3845: contig of 695 bp in length
3846 3945: gap of 100 bp
3946 4577: contig of 632 bp in length
4578 4677: gap of 100 bp
4678 5368: contig of 692 bp in length
5370 5469: gap of 100 bp
5470 6151: contig of 682 bp in length
6152 6251: gap of 100 bp
6252 6955: contig of 704 bp in length
6956 7055: gap of 100 bp
7056 7804: contig of 749 bp in length
7805 7904: gap of 100 bp
7905 8603: contig of 699 bp in length
8604 8703: gap of 100 bp
8704 9392: contig of 689 bp in length
9393 9492: gap of 100 bp
9493 10135: contig of 663 bp in length
10136 10255: gap of 100 bp
10256 10948: contig of 693 bp in length
10949 11048: gap of 100 bp
11049 11603: contig of 555 bp in length
11604 11703: gap of 100 bp
11704 12386: contig of 683 bp in length
12387 12486: gap of 100 bp

```

```

12487 13187: contig of 701 bp in length
13188 13287: gap of 100 bp
13288 13988: contig of 701 bp in length
13989 14088: gap of 100 bp
14089 14795: contig of 707 bp in length
14796 14895: gap of 100 bp
14896 15532: contig of 637 bp in length
15533 15632: gap of 100 bp
15633 16347: contig of 715 bp in length
16348 16447: gap of 100 bp
16448 17155: contig of 708 bp in length
17156 17255: gap of 100 bp
17256 17961: contig of 706 bp in length
18062 18723: contig of 662 bp in length
18724 18823: gap of 100 bp
18824 19476: contig of 653 bp in length
19477 19576: gap of 100 bp
19576 20274: contig of 698 bp in length
20275 20374: gap of 100 bp
20375 21074: contig of 700 bp in length
21075 21174: gap of 100 bp
21175 21867: contig of 693 bp in length
21868 21967: gap of 100 bp
21968 22680: contig of 713 bp in length
22681 22780: gap of 100 bp
22781 23495: contig of 715 bp in length
23496 23595: gap of 100 bp
23596 24227: contig of 632 bp in length
24228 24327: gap of 100 bp
24328 25032: contig of 705 bp in length
25033 25132: gap of 100 bp
25133 25838: contig of 706 bp in length
25839 25938: gap of 100 bp
25939 26651: contig of 713 bp in length
26652 26751: gap of 100 bp
26752 27431: contig of 680 bp in length
27432 28332: contig of 701 bp in length
28333 28332: gap of 100 bp
28333 29032: contig of 700 bp in length
29033 29132: gap of 100 bp
29133 29638: contig of 706 bp in length
29639 29938: gap of 100 bp
29939 30653: contig of 715 bp in length
30654 30753: gap of 100 bp
30754 31466: contig of 713 bp in length
31467 31566: gap of 100 bp
31567 32272: contig of 706 bp in length
32273 32372: gap of 100 bp
32373 33009: contig of 637 bp in length
33010 33109: gap of 100 bp
33110 33827: contig of 718 bp in length
33828 33927: gap of 100 bp
33928 34646: contig of 719 bp in length
34647 34746: gap of 100 bp
34747 35451: contig of 705 bp in length
35452 35551: gap of 100 bp
35552 36244: contig of 693 bp in length
36245 36344: gap of 100 bp
36345 36993: contig of 649 bp in length
36994 37093: gap of 100 bp
37094 37777: contig of 684 bp in length
37778 37877: gap of 100 bp
37878 38587: contig of 710 bp in length
38588 38687: gap of 100 bp
38688 39375: contig of 688 bp in length
39376 39475: gap of 100 bp
39476 40179: contig of 704 bp in length
40180 40279: gap of 100 bp
40280 40986: contig of 707 bp in length
40987 41086: gap of 100 bp
41087 41749: contig of 663 bp in length

```

```

41750 41849: gap of 100 bp
41850 42551: contig of 702 bp in length
42552 42652: gap of 100 bp
42653 43363: contig of 712 bp in length
43364 43463: gap of 100 bp
43464 44144: contig of 681 bp in length
44145 44244: gap of 100 bp
44245 44939: contig of 695 bp in length
44940 45039: gap of 100 bp
45040 45740: contig of 701 bp in length
45741 45840: gap of 100 bp
45841 46513: contig of 673 bp in length
46514 46613: gap of 100 bp
46614 47307: contig of 694 bp in length
47308 47407: gap of 100 bp
47408 48097: contig of 690 bp in length
48098 48197: gap of 100 bp
48198 48884: contig of 687 bp in length
48885 48984: gap of 100 bp
48985 49700: contig of 716 bp in length
49701 49800: gap of 100 bp
49801 50484: contig of 684 bp in length
50485 50584: gap of 100 bp
50585 51288: contig of 704 bp in length
51289 51388: gap of 100 bp
51389 52107: contig of 719 bp in length
52108 52207: gap of 100 bp
52208 52809: contig of 702 bp in length
52809 53009: gap of 100 bp
53010 53665: contig of 656 bp in length
53666 53765: gap of 100 bp
53766 54444: contig of 679 bp in length

```

```

Query Match      80.0%; Score 19.2; DB 2; Length 57662;
Beet Local Similarity 87.5%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 TGTGGATTCAGTCATCAACA 24
Db 1359 TGTGGATTCAGTCATCA 1382

```

```

RESULT 14
AL356577
LOCUS Homo sapiens chromosome 11 clone RPI-6104.
DEFINITION Homo sapiens
ACCESSION AL356577
VERSION AL356577.5 GI:11991384
KEYWORDS HTG; HTGS PHASE2; HTGS_CANCELED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Martin, S.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Dec 24, 2000 this sequence version replaced gi:11931890.
COMMENT ----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: d06104
----- Summary Statistics
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 77% of reads
Chemistry: Dye-terminator ET-amersham; 22% of reads
Consensus
quality: 84555 bases at least Q40

```

```

*-----*
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*-----*
Consensus quality: 84806 bases at least Q30
Consensus quality: 84862 bases at least Q20
Insert size: 85378; sum-of-configs
Insert size: 98007; 13.5% error; agarose-fp
Quality coverage: 8.00x in Q20 bases; sum-of-configs Quality
coverage: 7.06x in Q20 bases; agarose-fp

```

FEATURES	
source	location/Qualifiers
	1..85378
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="11"
	/clone="RP1-6104"
	/clone_1lb="RP1-1"
misc_feature	1..85378
	/note="assembly_fragment:01465"
ORIGIN	

Query Match	80.0%	Score 19.2	DB 2	Length 85378
Best Local Similarity	87.5%	Pred. No. 35		
Matches 21; Conservative	0	Mismatches 3	Indels 0	Gaps 0

RESULT	15
HSDJ247C2	
LOCUS	HSDJ247C2
DEFINITION	HSDJ247C2 98360 bp DNA linear PRI 03-AUG-2001
ACCESSION	Human DNA sequence from clone RPL-247C2 on chromosome 11p13
VERSION	AL049713
KEYWORDS	AL049713.20 GI:10334639 HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE	1 (bases 1 to 98360)
JOURNAL	Wallis,J. Direct Submission Submitted (02-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Sep 27, 2000 this sequence version replaced gi:606528.
COMMENT	

one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. Rp1-247C2 is from the library RpC1-1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2. (<http://www.chori.org/bacpac/home.htm>)

```

FEATURES
SOURCE
Location/Qualifiers
1. .98360
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="p13"
/clone="RP1-247C2"
/clone_1fb="RPC1-1"
complement(627..983)
/notes="match: STS: Em:HS198YB10"
804..829
/notes="13 copies 2 mer gt 100% conserved"
3773..3810
/notes="L1PAll repeat: matches 2154..2191 of consensus"
3956..3992
/notes="L1PAll repeat: matches 2180..2215 of consensus"
4265..5286
/notes="L1PAll repeat: matches 2215..3216 of consensus"
5366..8453
/notes="L1PAll repeat: matches 3216..6162 of consensus"
10697..11906
/notes="L1MEC repeat: matches 2142..1979 of consensus"
11202..11493
/notes="L1MEC repeat: matches 1978..2279 of consensus"
13794..13857
/notes="33 copies 2 mer cc 60% conserved"
14618..15078
/notes="L1R2 repeat: matches 1..449 of consensus"
complement(14642..15322)
/notes="match: GSS: Em:B04935"
15910..15911
/notes="single clone region. Assembly confirmed by
restriction digest data."
15930..15955
/notes="13 copies 2 mer tg 92% conserved"
16984..17077
/notes="MT1A2 repeat: matches 1..92 of consensus"
17380..17654
/notes="MT1A2 repeat: matches 92..374 of consensus"
complement(17710..18372)
/notes="match: GSS: Em:AQ485260"
18444..18750
/notes="match: GSS: Em:AQ54589"
complement(20477..21170)
/notes="match: GSS: Em:AQ48664"
complement(20761..21177)
/notes="match: GSS: Em:AQ66925"
complement(25626..26158)
/notes="match: GSS: Em:AQ50538"
25991..26104
/notes="match: GSS: Em:A239320"
26106..26561
/notes="match: GSS: Em:A239320"
29550..29636
/notes="L1MA8 repeat: matches 6175..6268 of consensus"
33334..33473
/notes="L1R23 repeat: matches 56..98 of consensus"
complement(34537..34987)
/notes="match: GSS: Em:AQ877394"
35999..36384
/notes="L1P3 repeat: matches 5743..6147 of consensus"
37497..37542
/notes="23 copies 2 mer tg 100% conserved"
38577..38776
/notes="match: STS: Em:G45608"
38617..38761

```

```

misc_feature      /note="match: STS: Em:A0048503"
                  38620..38722
                  /note="match: STS: Em:A0046429"
misc_feature      complement(38660..38777)
                  /note="match: GSS: Em:A2398168"
repeat_region     38685..38746
                  /note="11 copies 2 mer gc 91% conserved"
repeat_region     38956..39215
                  /note="Charliele repeat: matches 28..281 of consensus"
repeat_region     40351..40840
                  /note="Charliele repeat: matches 277..769 of consensus"
repeat_region     42053..42372
                  /note="Charliele repeat: matches 778..1137 of consensus"
misc_feature      44071..44491
                  /note="match: GSS: Em:A0112832"
misc_feature      complement(45407..45624)
                  /note="match: GSS: Em:A0247033"
repeat_region     53815..53904
                  /note="Charlie2 repeat: matches 3272..3370 of consensus"
repeat_region     54009..54235
                  /note="Charlie2 repeat: matches 2888..3117 of consensus"
repeat_region     56236..57009
                  /note="L1PA2 repeat: matches 1..776 of consensus"
repeat_region     57005..62248
                  /note="L1PA2 repeat: matches 900..6144 of consensus"
repeat_region     63508..63541
                  /note="17 copies 2 mer aa 82% conserved"
repeat_region     63586..63668
                  /note="L1MD repeat: matches 738..819 of consensus"
repeat_region     63917..64021
                  /note="L1MD repeat: matches 984..1089 of consensus"
misc_feature      complement(64074..64543)
                  /note="match: GSS: Em:A0885350"
misc_feature      complement(66096..66577)
                  /note="match: GSS: Em:A0611651"
misc_feature      69990..70342
                  /note="match: STS: Em:G21603"
misc_feature      complement(76094..76635)
                  /note="match: GSS: Em:A0277578"
misc_feature      complement(76184..76636)
                  /note="match: GSS: Em:A0812553"
misc_feature      complement(76460..76628)
                  /note="match: GSS: Em:A0140871"
misc_feature      complement(88581..89052)
                  /note="match: GSS: Em:A0201907"
misc_feature      89057..89594
                  /note="match: GSS: Em:A0543464"
repeat_region     91625..91990
                  /note="THE1C repeat: matches 1..371 of consensus"
repeat_region     92079..92237
                  /note="L1PA7 repeat: matches 5949..6125 of consensus"

```

ORIGIN

```

Query Match      80.0%; Score 19.2; DB 9; Length 98360;
Best Local Similarity 87.5%; Pred. No. 35;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Oy      1 TCTGGGATCTCAGTCATCAACA 24
          |||||
Db      36062 TCTGGGATTTAGTCATCAACA 36085

```

Search completed: October 15, 2005, 01:33:41
 Job time : 104.037 secs

This Page Blank (uspto)